THE UNIVERSITY OF ALBERTA

THE EFFECT OF SIMILARITY ON LEARNING:

AN EXPLICATION OF SIMILARITY WITH

APPLICATIONS TO LEARNING

by

© MARIA FLAK

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF EDUCATION

DEPARTMENT OF SECONDARY EDUCATION

EDMONTON, ALBERTA

SPRING, 1971

1971
44

This thesis is respectfully dedicated

to my father

and

to the memory of my mother.

ABSTRACT

The primary purpose of this study was to investigate the effect of similarity on learning.

An explication of the phenomenon of similarity led to the conclusion that similarity, being a relationship between two or more stimulus objects, has a potential for organization. It was then argued that if the subjects do not primarily learn by rote but look for ways to organize the material that they are to learn and if organization facilitates learning, then similarity should facilitate learning. However, an examination of the literature indicated that this was only true under certain conditions with meaningful materials. The effect of formal similarity was primarily negative.

Some of these effects were explained in terms of an interaction between the variables of similarity and learning conditions. However, it was argued, even under the most favourable learning conditions it would be unlikely that high formal similarity would lead to facilitation as long as the organizational potential was low. The formal similarity definition as it was stated did not take into account the spatial relationships among the elements of a stimulus object, nor did it consider their sequential constraints. As a result it was possible to construct two lists equal in formal similarity but differing in organizational potential. A new definition of similarity for nonsense syllables was formulated in terms of the three parameters of organizational potential so that similarity and organizational potential would vary directly. It was hypothesized that

performance on the list high in organizational potential (high similarity) would be superior to the performance on the list low in organizational potential (low similarity).

The effect of similarity on learning was investigated by means of a free recall experiment. Three lists equal in formal similarity but differing in terms of the new similarity definition were constructed. A fourth list, designated the no-similarity list, consisted of twelve different items. All trigrams were presented to the subjects simultaneously on the first trial. Each learning trial consisted of a two minute study period followed by a one minute recall period. After five, ten, or fifteen trials the subjects checked their last written recall. If all twelve items were correct they went on to the Questionnaire, if not, they continued for five more trials up to a maximum of fifteen trials.

The major findings supported the hypothesis that similarity (defined in terms of organizational potential) facilitates learning under free recall conditions. Therefore, a high degree of formal similarity does not necessarily hinder learning. The latter effect is only obtained if high formal similarity is combined with low organizational potential. Under conditions of high organizational potential formal similarity was found to facilitate learning. Since the new definition of similarity for nonsense syllables was stated in terms of organizational potential and varied directly with it, the results may be restated as follows: Performance varies directly with similarity (new).

The study also found that performance on the no-similarity lists was superior to that on the low similarity lists but inferior to that on the high similarity lists (although the latter differences were not statistically significant). These results were interpreted in terms of the three parameter model of organizational potential. It was argued that

since the high, low, and no-similarity lists could be placed at the three vertices of the organizational potential space there was no reason to expect performance to decrease linearly from high similarity to no-similarity.

ACKNOWLEDGMENTS

The author is deeply indebted to many people who contributed in so many ways to make this study possible. First of all, special thanks are extended to Dr. D. W. R. Wilson, committee chairman, for his wise counsel and editorial corrections of the manuscript, to Dr.H. Kass for her many valuable suggestions and encouragement, and to Dr. W. N. Runquist for his penetrating criticism which helped to deepen a number of insights. Thanks are also extended to Dr. M. A. Nay for his guidance throughout the initial and planning stages of the thesis.

The co-operation of students, teachers, and administrative staff of the following schools was much appreciated: Leduc Senior High School, Leduc; St. Albert High School, St. Albert; and Memorial Composite High School, Stony Plain, Alberta.

Thanks are also extended to Dan Precht and Kyung Bay for their assistance in setting up the statistical analysis of the data.

Special recognition goes out to Oksana Dexter for the many long and late hours that she contributed as research assistant.

Many persons, to numerous to mention, assisted in the preparation of data and the manuscript. To all, a sincere thanks.

Finally, and most important of all, a very special thanks to my father whose contribution extended from encouragement, through sacrifice, to constant and firm moral support.

TABLE OF CONTENTS

Appendixes

I. Tables

J. Analysis of Variance for Sex x Lists x Order

K. Orthogonal Comparisons

LIST OF TABLES

LIST OF FIGURES

Figure Page

Chapter 1

INTRODUCTION

The phenomenon of similarity has interested thinkers for over 2000 years. During that time it has been considered a principle of association (James, 1891), and a principle of perceptual organization (Koffka, 1935). Plato discussed the related concept of 'identity' in Parmenides, and Hume (1748) wrote about the similarity between ideas. With the growth of knowledge and science it has taken on a new significance. As Noble (1957) pointed out, problems concerning similarity can be encountered in such diverse areas as comparative biology, anthropology, and genetics; business and industry, particularly in cases such as patent, copyright or trade-mark litigation; and in education where researchers have stressed its role in concept formation, language development, reasoning, thinking and learning. The last, however, is the most important since it is the foundation on which the others have been built.

The role of similarity in learning takes on an added significance in science teaching since science as a body of knowledge and as a process, is highly dependent on similarity. Therefore, if the teaching of science is to reflect the process and nature of science, a clear understanding of similarity and its role in the process of scientific investigation and theory construction is necessary. Furthermore, since science deals with the physical universe and similarity is a definite and ubiquitous part of that universe, science is

also intrinsically concerned with similarity in nature at all levels of research and theory building. However, before the relationship between similarity and creative scientific endeavour can be understood and thus conveyed to the students, it is necessary to have a clear idea of the nature of each component in the relationship. Since it is no longer regarded as satisfactory to teach science only as a body of knowledge but rather as a process, it is no longer possible to teach science as just a systematic technique of experimental verification. Science teaching has to reflect the excitement of creative experience, and it is here that similarity plays a paramount role. However, exactly how creativity occurs and what its relationship is to similarity is not yet fully understood. Nevertheless, if the teachers are to structure meaningful scientific learning experiences for the students the relationship between similarity and creativity <u>has</u> to be understood and this can be done by first trying to understand <u>each</u> phenomenon separately.

Similarity, however, has a second role in science education which is common to all other subjects taught in the school, namely its role as a learning tool. Bruner (1959) indicated that learning consists of being able to identify recurrent regularity in the environment, indentification of which requires the construction of a model. Thus superior learning would depend on how well a given student was able to identify regularity. However, since regularity is a pattern created by the relationship between similarities and differences, successful learning should depend on the ability to perceive the similarities, that is, to abstract them from the 'ground'. The perception of the similarities among ideas, concepts, and theories can

lead to a structuring and organization of these into meaningful and useful wholes. The knowledge of many unrelated 'facts' is not only difficult to attain but is relatively unprofitable in the age of the computer library. On the other hand, the formation of a useful and comprehensive understanding of science, its processes, and its relationship to the world is a desirable goal.

Structuring, however, need not occur along the predetermined lines of organization available in a given discipline. In fact, the usefulness of an integrated perception of the subject depends on perceiving the pattern of similarities across different subject areas. Thus the perceived similarities among procedures or problem solving behaviors could be codified into models of thinking. In other words, there are two levels of similarities. At the lowest level similarities among the components of the subject matter are perceived and used to structure the subject. At the higher level, similarities among procedures and creative behavior can be used to formulate a model of thinking and learning. The tacit assumption underlying this interpretation of Bruner's statement is that similarity and the complementary concept of difference are the key elements in the thinking process. Consequently, by understanding similarity as a phenomenon and its use in the learning situation, a contribution could be made to the eventual development of a theory of learning and thinking.

The general importance of similarity in relation to learning has long been recognized in verbal learning studies and consonant with this, the phenomenon of similarity has been studied in many different ways. The result has been the production of a vast literature proposing a variety of explanations, supporting diverse approaches, using

different definitions, and reporting what seem to be conflicting experimental results.

An examination of the literature indicated that the different ways in which similarity has been studied can be grouped into four basic approaches which may be called theoretical, empirical, assumptive, and explanatory-definitive. The assumptive approach attempts to determine or measure the degree of similarity; however, the technique seems to rest on an assumed relationship between some overt action by a S and the degree of similarity. Thus, for example, the latency measure of similarity seems to assume that the longer it takes the Ss to verbalize one way in which a pair of stimuli are similar, the less similar the pair. Two other measures of similarity are the 'production of similar criteria' and the 'associative response'.[1] However, since in each case the relationship is only assumed and not proven, the three measures of similarity seem to be wide open to theoretical arguments of validity.

Traditionally similarity has been associated with the study of transfer, stimulus-generalization, interference, and retroactive inhibition (McGeoch and Irion, 1952; Osgood, 1953; Underwood, 1966). However, it is not at all certain if these phenomena are to define similarity or if similarity is to be used to explain the phenomena; hence the name: explanatory-definitive. Noble (1957) has this to say about stimulus generalization, for example:

> Much has been written about certain supposed relationships between stimulus generalization and similarity. Some assert the former to be a function of the latter, others assert the converse, and there are still others who deny that any relationship whatever exists. (p. 35).

---

[1]See Muir (1968) for a more detailed description of what has

However, it seems that both ways of looking at the relationship are equally valid as long as it is understood that <u>an independent conception, definition, and measure of similarity is also needed</u>, that is, in order to avoid circularity. If, for example, the above mentioned phenomena are to be explained in terms of similarity then surely implicit in the explanation is some conception or definition of similarity, regardless of whether it is stated or not. Furthermore, the relationship between the pehnomena and similarity has to be proven if theoretical arguments about the adequacy of the explanation are to be avoided. However, any attempt to prove the relationship would require a specification of a definition and measure of similarity independent of the phenomenon that it was trying to explain. On the other hand, if similarity is operationally defined in terms of any one of these phenomena then they, in fact, serve as measures of similarity. However, it seems that even here <u>an independent measure and definition of similarity is needed</u>. Torgerson (1958) discussing the relationship between operational definitions and the corresponding theoretical terms or concepts has presented a model which shows the relationship between the two types of constructs as well as what happens to the process of verification or rejection of a theory in such cases. In the model there are

> ...two sets of constructs: those on the right, lying very close to Nature, interconnected with one another, and each with a rule of correspondence connecting it to the observable data; and those on the left, also interconnected with one another but with no operational definitions (p. 6).

Furthermore, there are some dotted lines connecting the two groups of

---

here been called the 'assumptive' approach.

constructs. The set of constructs on the left "...might represent the 'socially or behaviorally important' concepts that a theoretically minded person might have in mind when he begins to work on constructing a theory..." (p. 6). Torgerson lists a number of such concepts as, for example, socioeconomic status, tension, intelligence, attitude, etc., and it seems that similarity might be added to this list. The connections between the concepts in the model indicate theoretical relationships which are perhaps most often "...expressed verbally rather than mathematically, and with a good deal less rigor" (p. 6). On the other hand, "...the set of constructs on the right...all possess operational definitions - rules of correspondence relating them to observable data. Very often they have the same names attached to them as those previously discussed" (p. 6). However, Torgerson cautions using intelligence as an example:

> But it is important to note that this operationally defined intelligence...is not universally agreed to be the same thing as the theoretically defined term having the same name. Hence the dotted lines. In these sciences [one often has] the concept as one thing _and_ a measure of it as another. The operationally defined construct on the right is an _indicant_...or an index of the equivalent on the left. At best, the two are _presumed_ to be monotonically related to each other. At worst, merely a positive correlation of unknown magnitude is _presumed_ to exist. A dotted line then stands for a presumed relationship between the construct with operational meaning and the equivalent concepts which possesses a wealth of constitutive meaning (p. 7).

Consequently, if a theory is tested and disagreement found, the researcher may

> ...conclude that the theory is incorrect, or he may equally well conclude that the opertionally defined variables are, after all, not really measures of the corresponding constitutively defined variables...(p. 7).

As far as similarity is concerned this model clearly indicates that in addition to an operational definition (such as transfer, for

example, or even one of the 'assumptive' definitions) there is a need for a <u>well defined theoretical conception of similarity</u>. It seems that there are, on one side, the concepts of theoretical interest which tend to lack empirical meaning, and on the other hand, the corresponding concepts with precise meanings but often lacking theoretical import. According to Torgerson, "One of the great problems in the development of a science is the discovery or invention of constructs that have, or are likely to have, both" (p. 8).

The theoretical concepts such as similarity, for example, very often have much common-sense meaning attached to them, but this is not specified precisely; and thus the terms are somewhat vague and often complex. "Before a satisfactory state of affairs is reached, it is necessary somehow to transform these inexact, complex concepts into exact ones which can be specified precisely" (Torgerson, 1958, p. 8). In the words of Carnap (1950) these theoretical concepts have to be "explicated". This process is defined as "...the transformation of an inexact prescientific concept, the <u>explicandum</u>, into a new exact concept, the <u>explicatum</u>..." (Carnap, 1950, p. 3). Although the process seems similar to analysis and has the same motive, namely, that of "...supplanting a relatively vague idea by a more precise one..." (Langford, 1943); nevertheless, it is broader in scope.

> The explicatum...is in many cases the result of an analysis of the explicandum....in other cases, however, it deviates deliberately from the explicandum but still takes its place in some way... (Carnap, 1950, p. 3).

Carnap (pp. 5-6) illustrates this with the scientific and non-scientific term 'fish'. In every-day usage 'fish' refers to 'animals living in water' and this includes whales and seals. The scientific term,

however, is much narrower and refers only to those water animals which are cold-blooded vertebrates and have gills throughout most of their life. This excluded whales and seals; however, this does not mean that a correction was made in knowledge but rather that a change was made in the rules of languages. In scientific circles the old term (explicandum) was no longer necessary and had been replaced with a much more precise explicatum.

The primary criterion for the evaluation of the explicatum is whether is is 'fruitful', that it, "...useful for the formulation of many universal statements..." (Carnap, p. 7). As such, the solution of an explication cannot be judged as being 'right or wrong'; but rather, as being 'more satisfactory than another one'.

Similarity is a concept of great theoretical interest; however, relatively few recent attempts have been made to explicate it (Goldmeier, 1937; Landahl, 1945; Attneave, 1950; Noble, 1957; and Wallach, 1958). Nevertheless, since the time of Plato the concept of similarity has undergone sufficient explication to yield two models of similarity: the 'common elements' and the 'dimensional' model. However, both models have certain weaknesses, and with certain materials they are unable to give satisfactory explanations. Furthermore, an examination of the literature with respect to the effects of similarity on learning has shown that many of the contradictory experimental results can be traced to the 'formal similarity' definition which, in turn, can be traced to the 'common elements' model of similarity. Thus, it seemed that further explication of the similarity concept was needed, particularly if the conflicting effects of similarity (formal and meaningful) were to be understood and

resolved. Consequently, the problem selected for this thesis is the following: An explication of the theoretical concept of similarity and the resolution of the seemingly contradictory effects of similarity on learning.

However, any examination of similarity would not be complete without mentioning another approach to explication. Whereas the former attempts at explication may be termed the theoretical approach to similarity, the latter may be referred to as the empirical approach because it tries to explicate similarity by starting with observable data. The two approaches yield different models of similarity and although only the theoretical one is of interest, nevertheless, any comprehensive treatment of similarity must include both and show the relationship between the two.

Chapter 2

THE PROBLEM

A preliminary examination of the literature indicated that although similarity has been studied in many different ways the attempts at explication fall into the two expected categories: the primarily theoretical, and the predominantly empirical mode of attack. The theoretical approach can be further subdivided into the 'common elements' model and into the 'dimensional' model. The former is of particular interest because it gave rise to the 'formal similarity' definition which is now universally used in verbal learning experiments; and it is here that the 'similarity paradox' occurs, assuming, of course, that similarity is one phenomenon and therefore should have a consistent effect on learning under comparable conditions.[1] Experiments which examine the effect of meaningful and formal similarity on learning seem to come to the untenable and paradoxical conclusion that while meaningful similarity facilitates learning or at least has no significant effect on it, formal similarity either has no significant effect or hinders learning. A further examination of the experimental data in the literature suggested that the major source of the paradox might be the 'formal similarity' definition. Consequently, in order to resolve the paradox, it was necessary to examine not only the effect

---

[1]For an explanation and definition of terms used in the text, see Appendix A.

of similarity on learning but also the theoretical approach to similarity, particularly the 'common elements' explanation or model. However, a thorough understanding of the problem could not be obtained without also examining the empirical approach to similarity. While the theoretical approach attempts to answer the question 'What makes things seem similar?', the empirical approach tries to predict as accurately as possible future similarity judgements. Both approaches have a similarity model as a goal but while the former arrives at it after analysis of the stimulus objects (StOs) that seem similar, the latter starts with subjects' (Ss') similarity judgements and resolves these into dimensions which when recombined would yeild the Ss' judgements.

The following sections will now briefly summarize each approach and indicate the major areas that are still in need of further explication. One section will also be devoted exclusively to an examination of the literature dealing with the effects of similarity on learning.

The Theoretical Approach to Similarity.

Attempts to answer the question: 'What makes things seem similar?' have resulted in two major approaches. They may be called the 'common elements' and the 'dimensional' explanations of similarity.

The first states that the similarity between objects and events varies directly with the number of elements held in common. The 'elements' may be the 'simple ideas' of Hume, the 'common attributes' of James, or the 'common letters' in nonsense syllables which have so often been used in recent learning experiments. Thus, the 'elements' are conceptualized as being discrete entities and similarity is determined by a counting procedure.

The 'dimensional' approach defines similarity in terms of proximities on the various dimensions that the objects exhibit. This presupposes some means of measuring the distance on a given dimension and a way of meaningfully combining the resultant set of proximities into one common index of similarity.

Each approach has its supporters and critics; however, not all the problems have been solved. For example, in a given set of stimulus objects (StOs) there may be some that do not have any dimensions or elements in common. This makes the similarity relation non-transitive (Noble, 1957). As far as the 'common elements' approach is concerned this is acceptable. However, for the dimensional approach it poses grave difficulties since dimensions are continua and indicate similarity by proximity on a given extent rather than on the basis of <u>the presence or absence of the whole dimension</u>. Attneave (1950) feels that this whole area is in "...considerable need of clarification" (p. 554).

On the other hand, the 'attributive' approach which is a very important type of 'common element' explanation faces grave difficulties in dealing with StOs that possess attributes which refer only to a part of the StO. Most attributes refer to the whole StO; for example, the whole moon is round not just a part of it. However, there are StOs which are more complex, so that a given attribute can refer only to one of its parts. Thus, chairs, for example, can exhibit characteristics of roundness and angularity depending on which part one is examining. When this happens attributes form spatial relationships. Nevertheless, this is one problem which has been mostly ignored by theoreticians and researchers alike. This is unfortunate because the

definition of formal similarity used universally in learning experiments is taken directly from the common elements interpretation of similarity and therefore has the same weakness as the parent definition. As the discussion in Chapter 5 will show, many of the seemingly contradictory effects of similarity on learning may be traced directly to this weakness in the parent definition. Furthermore, it seems that the whole problem of the spatial relationships between common attributes or common elements might never have occured if the relationship between similarity and identity had been considered in greater detail. For example, _if_ identity is considered as 100% similarity, then 100% common elements should yield identity. However, this is not the case between the two StOs of the pair ABC-BCA. Thus, obviously, something more than the NUMBER of common elements must be considered in determining similarity, namely the spatial co-ordinates of the common elements. This aspect will also be discussed in great detail in Chapter 3 and particularly in Chapter 5 where a new definition of similarity for nonsense StOs will be proposed.

The consideration of 'identity' brings up the very interesting philosophical question of whether two StOs can ever really be identical. Plato answered this question in Parmenides some 2000 years ago by pointing out that no two objects can possibly be identical since they cannot occupy the same space-time co-ordinates without being one and the same object. This conclusion has very significant implications for the concept of similarity because not only does it point out the importance of spatial co-ordinates as a characteristic but also because it points out that there may be a need to distinguish between two types of stimulus characteristics: those which strictly pertain to

the StO, and those which relate the StO to the environment.

Hume (1748), writing on the similarity among ideas, reduced these to sensory impressions and then determined similarity only on the basis of the physical characteristics that they had in common. This enabled him to explain the similarity between complex ideas since this was the number of simple ideas held in common, but it did not enable him to explain how it was that two simple ideas such as blue and green were similar. Wallach (1958) suggested that Hume would have solved the problem if he had admitted those characteristics which related the StO to the environment. For example, blue and green are similar because they are the colours of water, of grass in the country, and summer landscape colours, etc. This, however, does not solve the problem; in fact, it makes it unmanageable. For example, one such environmental characteristic would be 'use'; however, this is one of the most variable attributes often used in tests of creativity and is only limited in scope by the _Ss'_ imagination. Thus, for example, a stone and a paper clip would have to be considered as similar because both can be _used_ to make necklaces! Is this to be considered as a valid similarity? If not, where is the line to be drawn, and by whom are its criteria to be set? That some limit is needed is evident because if none is set then _everything will be similar to everything else_ and similarity will depend on all and any idiosyncratic associations. Although this _is_ one way of looking at similarity it does not seem to be the most fruitful approach since, under such circumstances, similarity would, in effect, become indeterminable. However, if Wallach's solution is rejected then Hume's original problem still remains.

These, then, are some of the major problems that seem in need

of explication in the theoretical approach to similarity.

The Empirical Approach to Similarity.

The empirical mode of attack grew out of psychophysics which was developed by Gustav Theodor Fechner (1801-1887). He attempted to find 'functional relations of dependency between body and mind'. These ideas were later developed by Louis Leon Thurstone (1887-1955) and when appropriate statistical techniques were formulated by Young and Householder (1938), the method of multidimensional scaling was born. This technique approaches similarity from the observer's point of view. The experimenter obtains judgements of similarity from Ss, converts these into proximities and then applies statistical treatments very much like factor analysis. The result is a number of dimensions which presumably the Ss used in arriving at their similarity judgements.

The multidimensional technique of Young and Householder was specifically developed on the request of Gulliksen and Richardson for the purpose of scaling psychological judgements in general. It was first applied by Richardson (1938) to the judgements of similarity. In this study the Ss were presented with three StOs at a time and had to judge which of the two StOs were most similar and which two were most different. From this the proportion of times that a particular StO was judged closer to one of each possible pair of StOs was obtained. This in turn, was converted into a set of 'differences in distance between the StOs'. According to McCullough (1957) the procedure was essentially an extension of Thurston's technique using the normal probability curve for converting proportions into dis-

tances. At this point in the process it was necessary to have some means of transforming the 'distances' between the StOs into measures along some set of dimensions or axes using the least number of these necessary for a full description of the system. In other words, the StOs were seen as points in some sort of space, their location relative to each other (their mutual distances) being determined by the similarity judgements. That is, "...similarity was considered to be the complement of distance in a space of one kind or another" (Torgerson, 1965, p. 379). If the similarity judgements were to be resolved into the contributing dimensions which presumably the Ss used in arriving at their judgements, then a mathematical technique was needed which could determine not only the minimum number of dimensions necessary to describe the system of points but also to yield the co-ordinates of all the points relative to this reference system.

> Ordinarily a set of points is specified by giving its co-ordinates in a suitable reference system; and the dimensionality of the set, the problem of approximating it by a lower dimensional set, etc., can be discussed in terms of these co-ordinates. It may be, however, that only the distances of the points from each other are known, and it is desired to give a similar discussion on this basis (Young and Householder, 1938, p. 19).

This is exactly what the technique devised by Young and Householder (1938) was designed to do. However, any mathematical technique is based on a number of assumptions and has some formal requirements which must be met if it is to work and give meaningful answers. Torgerson (1965) describing the subsequent development of this procedure by Gulliksen (1946), Messick and Abelson (1956), and Torgerson (1952, 1958) states that although this traditional procedure or

set of procedures 'knew what it was doing and what it needed', its formal requirements were too severe.

> It asked not only that the perceptual or cognitive structure of the set of stimuli be Euclidian in nature, but also, that observations on similarity of pairs of stimuli be linearly related to distances between points in the space (Torgerson, 1965, p. 379).

In other words, the use of this mathematical technique required that some assumptions be made as to the nature of similarity which could not always be made. In short, if the mathematical model was to give psychologically meaningful answers the same relationships would have to hold in the psychological system as in the mathematical model. If, however, the mathematical requirements were too restrictive or psychologically untenable then the model could only have a limited applicability.

The next development in multidimensional scaling came in 1962 with the advent of the first computerized program. It was spearheaded by Shepard (1962a & b) and his papers on the 'analysis of proximities'. Since then a number of researchers have proposed their own versions (see Torgerson, 1965, p. 380). These, however, deal mostly with different methods of defining and determining the various proximities which are to be used in the matrices, and of determining the location of the origin by judicious selection of the 'additive constant'.

The basic advantage of the new iterative procedure is that it does not require the assumption of a Euclidian metric but works equally well for a non-Euclidian one, and in fact, for any metric. This is a particularly significant improvement over the traditional multidimensional technique because there is no valid reason to assume a priori that the space should be Euclidian; and in fact, there is evidence that

it is in many instances non-Euclidian (Attneave, 1950; Shepard, 1964). The raw data with which the experimenter starts is a set of 'judgements' rather than physical measurements, and thus the dimensions that are identified by the multidimensional technique are described as '_psychological_ dimensions'. Hence there is much more reason to expect that the space will be non-Euclidian than Euclidian. However, this may to a large extent depend on the nature of the StOs.

The overall aim of the multidimensional technique is to _identify_ the _psychological_ dimensions that are presumably used by the _Ss_ in making the similarity judgements. There is, however, another empirical approach to similarity which attacks the problem from a slightly different direction. Assuming that the dimensions are _known_, one can try to determine how they would combine so as to give the impression of similarity. This is the method that Attneave (1950) has used. He developed what he called a 'method of graded dichotomies' for scaling similarity judgements (1949). In this procedure the _Ss_ are presented with pairs of stimuli which they are to judge as to degree of similarity. Usually the _Ss_ are provided with a seven point scale ranging from 'identical' to 'extremely different' on which they mark their judgements. In the graded dichotomies technique, similarity is also considered as the complement of distance; however, the space is considered to be non-Euclidian and what is referred to as the 'city block' metric is applied to the distances. In this metric the distance (similarity) between two StOs is equal to the sum of the projections along each axis or dimension. In the Euclidian metric this distance, of course, would be calculated by using the Pythagorean theorem. Torgerson (1965) discussing the merits of Euclidian versus

non-Euclidian metrics concluded that both might be applicable depending on the nature of the StOs or more specifically the nature of the constituting dimensions.

The most important difference between these two empirical techniques, however, is the way in which the dimensions are identified. The multidimensional procedure gives the number of dimensions and the location of the StOs in the space bounded by these dimensions. The experimenter then attempts to interpret the dimensions in such a way that they are psychologically and/or physically meaningful. The graded dichotomies technique, on the other hand, gives only the mutual distances between the StOs, and the experimenter has to arrive at the number and the nature of the dimensions by inspection of the StOs, usually prior to the experiment. Thus, for example, McCullough (1957) using simple geometrical figures (triangles), constructed three sets of StOs such that the triangles in each set varied a priori on what he referred to as 'one physical' dimension (p. 38). However, it is difficult to think of 'area' or 'shape' as being single physical dimensions. Nevertheless, since the object in this approach is to find by configuration what seemed, a priori, to be psychologically relevant dimensions, (that is, those which the Ss might have used in arriving at their judgements of similarity) any physical measures may be used if the experimenter feels that they might be highly correlated with the scaled similarity judgements. However, it seems that such measures on the StOs should not be referred to as 'single physical dimensions' but rather as 'physical measures' or as 'single psychological dimensions' if they should prove to be relevant in accounting for the variance in the similarity judgements. There is, of course,

no reason to assume that there must be a one-to-one correspondence between the dimensions in psychological space and physical space. Thus, for example, the Ss could very well be responding to an overall gestalt of 'bigness' or 'area' rather than to the length of the base in combination with the height of the triangles. The success in using the graded dichotomies technique depends, therefore, to a large extent on the intuitive understanding of what the Ss might use in judging the StOs. However, the multidimensional scaling technique faces exactly the same problem after the mathematical model has 'ground out' the answer. That is, although the mathematical technique will indicate the number of dimensions and the location of the StOs relative to these dimensions, the researcher has to interpret these dimensions in terms of psychologically relevant variables. Nevertheless, the researcher using the multidimensional approach has some advantage over the one using the graded dichotomies method because the mathematical technique of the former will locate the StOs in the space and thus indicate to the researcher which set of StOs have a high loading on one dimension and which StOs have a very low loading on the same dimension and thus by comparing these two sub-sets of StOs the researcher can get some idea as to what the dimension might represent. Again there is no reason to assume that for each physical dimension exhibited by the StOs there should be a corresponding psychological dimension. Furthermore, it cannot be assumed that even if all the physical dimensions exhibited by the StOs contributed to the similarity judgements that they would do so equally. For example, Chambliss (1957) had Ss judge for similarity very complex StOs (ten-sided nonsense polygons) which might be considered as physically multidimensional; however, he

found that most of the variance could be accounted for by one physical measure, and that the correlation coefficient could be raised only slightly (from 0.9467 to 0.9580) by using three other physical measures. He concluded that there was only one psychological dimension involved and that it could be approximated by one physical measure. The latter, however, could not really be regarded as a single physical dimension nor as a composite dimension serving as a resultant of all the other physical dimensions. Thus, it seems that one psychological dimension may be represented by a combination of some of the physical dimensions exhibited by the StOs.

In summary, the expirical approach to the explication of similarity starts with a set of judgements and hopes to end with a model of psychological similarity. Of the two main ways of achieving this goal, multidimensional scaling is by far the most comprehensive technique. It starts with a set of similarity judgements, converts these into inter-StO distances or indices of similarity, and then applies mathematical techniques similar to factor analysis which gives the dimensionality of the space and helps the researcher to interpret the resulting psychological dimensions. The graded dichotomies technique, on the other hand, is primarily a scaling method which yields only the inter-StO distances or indices. The dimensions are assumed *a priori* and the distances are used to determine the nature of the space, that is, to find out how the dimensions combine into the resultant index of similarity or scale value. Alternately, the experimenter may calculate correlation coefficients between various physical measures, dimensions, or combinations of dimensions and the scaled similarity judgements. In this way the graded dichotomies technique is

much closer to the traditional goal of psychophysics, that of finding the 'functional relations of dependency between body and mind'.

Discussion

The empirical and theoretical attempts at explication have approached the problem from opposite directions. The latter started with an analysis of the StOs into physical dimensions and attempted from this to determine why the StOs seemed similar to the Ss; the former, on the other hand, started with the Ss and their similarity judgements and worked backwards to the StOs but ended up with psychological dimensions instead of physical ones. Attneave's method (graded dichotomies) basically attempted to work from both ends and to relate the physical properties of the StOs to the psychological ones as indicated by the scaled similarity judgements. Whatever the approach the general aim was the greater understanding of the phenomenon of similarity. However, what seems to be needed at this time is some sort of synthesis or structuring, particularly if the paradoxical effects of similarity on learning are to be understood and the conflict resolved.

An examination of the theoretical and empirical approach to similarity suggests that they are in one sense antithetical. The former to a large extent deals with the physical dimensions of StOs, and the latter is exclusively concerned with either determining the psychological dimensions or in relating them to the physical ones.

This dichotomy has great significance for learning experiments where the definition of similarity that is presently being used is

mostly in terms of the physical characteristics of the StOs.[2] However, if physical dimensions and psychological dimensions do not correspond exactly then physical similarity is not the same as psychological similarity; and if the Ss respond to psychological similarity during learning while the experimenter varies the degree of physical similarity, what exactly is the relationship between 'similarity' and learning?

To answer this question it is necessary to consider the relationship between physical and psychological similarity. Since neither has a monopoly on the 'truth' both have to be examined critically. Any definition of similarity which the theoretical approach produces has to be examined as carefully as any theory in science. If the definition is found to be too narrow or unable to explain or subsume all the observable data, then it has to be replaced with a better definition which, in its turn, will later be replaced with a better one still. Philosophically, it can be argued, and with merit, that one can never truly know absolute reality but only approximate it with temporary conceptions. However, with repeated attempts it is hoped that ever better approximations may be obtained, and at the moment an attempt at some sort of structuring of similarity seems to be in order. There is a need to collate all the major approaches, to bring together

---

[2]Phrases such as 'physical characteristics' or 'physical similarity' may be somewhat misleading, particularly when the StOs are meaningful words. More appropriate terms to use might have been 'intrinsic characteristics' (Chapter 3) and 'objective similarity'. In this way common categories and denotative meaning could relate in a much more meaningful way with the name. In the same way any similarity arising out of conotative meaning, associative overlap (Deese, 1962), or any idiosyncratic associations might be much more appropriately designated by the label 'subjective' rather than by the term 'psychological' similarity. However, since the terms 'physical' and 'psychological' have gained some

all the relevant data, and to examine them critically for weaknesses, strengths, and possible inconsistencies. This is particularly important if the seemingly paradoxical effects of similarity on learning are to be understood and some explanation found.

The Effect of Similarity of Learning: A Paradox

The role of similarity in verbal learning has been studied primarily through free recall, serial learning, and paired-associate (P-A) experiments.

In a free learning task, a series of verbal units are presented to the Ss with instructions to learn the units in any order. On successive trials the presentation order is either constant or varied from trial to trial depending on what is being studied.

A serial learning task involves the presentation of the units to the Ss in a constant order and the Ss are required to learn them in the order of presentation by anticipating the next unit in the sequence. The fixed order is originally determined on a random basis.

A typical P-A task requires the Ss to learn pairs of verbal units. The left-hand member of the pair is called the stimulus term and the right-hand member the response term. The Ss must learn to be able to recall the correct response term when presented with the stimulus term alone. The units are usually presented in a different order on each trial.

In all three tasks the units are usually presented one-at-a-

---

measure of acceptance, the author will continue to use them but with the added qualification that the term 'physical characteristic' is to be interpreted in a broad sense (Chapter 3).

time on an instrument called a memory-drum (McGeoch and Irion, 1952). The average viewing time is two seconds per item or stimulus term. In a P-A task this two second presentation-anticipation phase is followed immediately by another two second interval in which the stimulus term is shown together with the correct response term. The material to be learned is either of a meaningful or meaningless nature. At times the P-A task combines both, that is, either the stimulus or the response term is meaningful while the other member of the pair is a nonsense syllable. The most common type of nonsense material used is the trigram which consists of three upper case letters. These are either all consonants (CCC) or of the form consonant-vowel-consonant (CVC).

Similarity among the trigrams of a list is determined by the commonality of letters and is almost universally referred to as 'formal similarity'. Underwood (1966) defines this type of similarity as follows:

> Formal similarity refers to the number of different letters used in making up a list; the fewer the number of different letters for a constant number of units, the higher the similarity (p. 476).

The quantitative nature of the definition lends itself well to a precise determination of the degree of similarity present in a list of given length. The same cannot be said for meaningful similarity.

There are two common types of meaningful similarity; one is based on synonymity among the words (icy, frigid, cold) and the other on common class membership (dog, cat, cow, pig).[3] Some refer to the

---

[3] A third type of meaningful similarity may be based on the number of common associations elicited by a pair of words (bird-wing, fast-slow, sky-blue, blue-yellow). For a discussion of associative meaning and associative overlap see Deese (1962).

latter as 'conceptual similarity'. In either case, however, it is impossible to calculate the exact degree of similarity. Consequently researchers prefer to use lists constructed from nonsense syllables.

In a P-A list there are at least four locations for similarity, the most common being similarity either among stimulus terms, or among response terms. Less common is similarity between the stimulus term and its corresponding response term, or similarity among all stimulus and response terms.

The latter P-A list is one of the most difficult to learn when it has high formal similarity. In one study Underwood and Richardson (1957) had 336 college students learn a list constructed from only nine consonants with each letter occurring one to four times per list. Although the list had only four pairs of CCC's they had to drop 135 Ss from the experiment, of which 60 were dropped for failure to learn the list in a period of 50 minutes or approximately 150 trials.

Another list which is extremely difficult to learn is the 'double function' list in which a given trigram serves both as a stimulus and as a response term (Umemoto and Hilgard, 1961). Thus if the letters A, B, C, D, E, F, G, and H are taken to represent eight different trigrams the following pairs might be formed: A-D, B-G, D-H, E-C, F-B, C-A, H-E, and G-F. The S is faced with the task of learning to respond D when presented with A, and to respond H when presented with D. Since both D's are one and the same trigram, there is no question of similarity but of identity. A modification of the list so as to study the effect of similarity can be achieved by pairing CVC's with synonymous adjectives as follows:

YAV-perfect
faultless-YAV
noonday-GID
GID-midday etc.

As can be expected, such a list is easier to learn than one where the identical CVC's and the similar adjectives are mismatched (Umemoto and Hilgard, 1961).

The more typical P-A experiments use nonsense syllables for both the stimulus and response terms. The results of these experiments are somewhat inconclusive since experimenters did not always use the same measures of acquisition. Underwood (1953a), using mean number of trials to criterion as the measure of acquisition, found that only similarity among <u>stimulus</u> terms had a detrimental effect on learning. The effect of similarity among response terms was statistically not significant but tended to suggest that intermediate similarity was facilitative. The effect of similarity on recall 24 hours after original mastery proved to be facilitative only for high stimulus similarity. However, this could be explained by the additional trials required for mastery.

A similar experiment was conducted by Newman and Buckhout (1962) except that the mean number of correct responses on trial #39 was used as the measure of acquisition. Contrary to Underwood's findings, the results showed that the inverse relationship existed between <u>response</u> similarity and learning; and that the effect of stimulus similarity was not statistically significant. These results are not at all surprising in view of the different measures of acquisition used.

In his study Underwood (1953a) also examined the mean number of overt errors per trial and found that: "Overt errors increased

directly with response similarity but not with stimulus similarity" (p. 141). A clear cut comparison cannot be made with the Newman and Buckhout study because Underwood used overt errors as a measure of acquisition rather than errors (overt errors plus omissions); an inverse relationship between the number correct and the number of overt errors cannot be assumed. A study by Rubin and Brown (1967), on the other hand, using the mean total errors for trials to criterion (overt plus omission errors) as a measure of acquisition, found that not only was there a direct relationship between response similarity and total errors but that the same relationship held for stimulus similarity.

One interesting factor that does emerge from Underwood's study is the strange relationship between the number of overt errors per trial and the number of trials to mastery when performance on stimulus similarity lists is compared with that on response similarity lists. If stimulus similarity lists require more learning trials to mastery and have fewer overt errors per trial than response similarity lists, then surely there is an inverse relationship between overt errors per trial and number of trials to mastery. Underwood (1953a) sums this up as follows: "Thus, while learning was more rapid [requiring fewer trials to mastery] with variation in response similarity than with variation in stimulus similarity, more [overt] errors per trial were made in the former instance" (p. 137).

These results were 'substantiated' by Levitt and Goss (1961), using the mean number of correct responses for 20 trials as a measure of acquisition. They found that "Significantly more correct responses were made during the 20 trials with variation in response members

(M = 102.9) than with variation in stimulus members (M = 94.8)", (p. 249). This can be taken as supporting Underwood's finding to some extent because trials to criterion and correct responses during 20 trials may be considered to vary directly. However, no comparison can be made between these findings and those using only overt errors as a measure of acquisition.

Experiments using meaningfully similar material are also inconclusive. A study (Underwood, 1953b), using adjectives for both stimulus and response terms, found that while variation in response similarity had no significant effect on the number of trials needed for mastery, variation in stimulus similarity did affect the learning rate, but in an unexpected way. As the similarity increased, trials to learn first increased and then decreased. Thus, lists with medium stimulus similarity were most difficult to learn and were followed in difficulty by high similarity and low similarity lists respectively. The relationship between overt errors and similarity was in agreement with Underwood's (1953a) previous findings that only response similarity affects the number of errors significantly. The effect of similarity on recall 24 hours later was not significant when the degree of learning was held constant.

However, when trigrams replace adjectives as stimulus terms, meaningful <u>response</u> similarity also affects acquisition (number of trials to criterion) in a negative way (Feldman and Underwood, 1957). Furthermore, there is a significant and direct relationship between errors and variation in response similarity and <u>also</u> between errors and stimulus similarity (formal).

Attempts to resolve these conflicting findings have led to

further research where the learning process is broken down into two parts: response learning, and formation of associations between stimuli and responses. The results from one such study (Underwood, Runquist, and Schulz, 1959) indicate that meaningful similarity among response terms facilitates the acquisition of the responses but hinders the learning of specific associations between a given nonsense stimulus and its corresponding meaningful response. Together the two effects appear as an inverse relationship between similarity and the mean number of correct responses per trial.

Horowitz (1962) eliminated the response learning phase by substituting a matching task for the standard anticipation technique. For some lists the stimulus term was a nonsense syllable while for other lists the response term was a trigram. In all cases the other term was a meaningful word. The results of this study indicated that variation in either formal stimulus or formal response similarity hindered the associative phase.

A somewhat different approach was taken by Umemoto (1962). The similarity among the stimulus terms and that among the response terms was held constant at a high level of similarity while that between the stimulus and response sides was varied. Thus for the low similarity list the stimulus terms were selected from one pool of highly similar adjectives and the response terms from another. For this condition there was no similarity between the two pools of adjectives. For the medium similarity list there was some similarity and for the high similarity lists the stimulus and the response terms were drawn from the same pool of adjectives. As can be expected, the results showed that the higher the similarity between the stimulus and response terms, the faster the learning as defined by trials to criterion.

This suggests that Ss are able to overcome the hindering effects of either stimulus or response similarity if provided with sufficiently high stimulus-response (S-R) similarity as in Umemoto's (1962) study. However, this seems to hold only for meaningful material. A high degree of formal similarity among all terms of a P-A list made learning almost impossible (op. cit. Underwood and Richardson, 1957).

On the basis of all the above findings it seems logical to hypothesize that a list consisting only of single meaningful rather than paired elements should be facilitated by inter-item similarity. This is consistent with the findings that response learning as such is facilitated by response similarity (Underwood, Runquist, and Schulz, 1959). Furthermore, a serial learning task requires that the Ss learn these items in a given order; therefore, it might be assumed that the terms in the sequence have to form some sort of associative chain. This should be facilitated by high S-R similarity even in the face of high list similarity. This is not contrary to the findings of Horowitz (1962) because he did not have any similarity between S-R terms only among them. Also, he studied formal similarity.

Nevertheless, contrary to expectation, serial learning was found to be either unaffected by meaningful similarity (Baddely, 1966), or to be hindered by it (Underwood and Goad, 1951). However, as Underwood (1953c) found, the difference between the low and high similarity lists had to be made very great indeed in order to produce any statistically significant retarding effects. Under such conditions a very important fact was discovered: Recall, 24 hours after learning, was vastly superior for the high similarity list (significant at the

0.005 level) even when the number of learning trials was held constant.

On the other hand, formal similarity does not seem to affect recall significantly (Underwood, 1952; Underwood and Richardson, 1958). However, it does produce a hindering effect on learning as far as trials to criterion or mean number of correct responses for 30 trials are concerned (Underwood, 1952; Underwood and Richardson, 1958 respectively). Nevertheless, contrary to the findings of the 1958 study the 1952 experiment indicated that formal similarity was not significantly related to the number of errors during learning.

A somewhat different experiment by McGourty (1940) showed that the usual U-shaped serial position curve could be altered to a W-shape by grouping the high similarity CCC's at the beginning and end of the list. Thus, the learning of the centrally placed items was facilitated and that of the extreme items was hindered.

Turning to free learning experiments, the findings are again contradictory. Although the effects of meaningful similarity are either facilitative (Ekstrand and Underwood, 1963; Bourne and Parker, 1964; Underwood, 1964, pp. 63-65) or at least not significant (Cofer, 1959); the effects of formal similarity are either not significant (Allen, Puff, and Weist, 1968; Newman, 1967) or negative (Horowitz, 1961; Carterette, 1963; Underwood, Ekstrand, and Keppel, 1964). Furthermore, Horowitz, Carterette, as well as Allen, Puff, and Weist reported that initially formal similarity facilitated recall and that it was only on later trials that the Ss learning the low similarity lists showed superior recall. Underwood, Ekstrand and Keppel (1964), however, found no such effect.

In summary therefore, it seems that on the whole

> Conceptual and meaningful similarity are positively related to free learning; that is, the higher the similarity, the more rapid the free learning ... [whereas] ... with low meaningful units, high similarity retards free learning ... (Underwood, 1966, pp. 477-478).

A Summary Statement of the Problem and Purpose

In brief the problem may be summarized in terms of the following questions:

1. Assuming that similarity is one phenomenon, why does meaningful and formal similarity affect learning in different ways under comparable learning conditions?

2. What are the relationships among the concepts of identity, similarity, and difference?

3. What roles should the StO-pertinent and environmental characteristics play in a model of similarity in order to ensure maximum usefulness of the model?

4. Is the internal organization of the elements in a StO, and particularly the spatial relationship among the elements composing the StO, a crucial factor in determining the fruitfulness of a model of similarity?

5. How can mutually exclusive dimensions be explained by a dimensional model, or is the model completely inadequate in this respect?

6. What relationship, if any, if there between the dimensional and common element model?

7. What relationship, if any, is there between the empirical and theoretical models of similarity?

The purpose of the study, therefore, was:

1. to answer questions two through seven, thereby explicating the concept of similarity to some extent;

2. to relate the explicated concept of similarity to learning by exploring the variables relevant to this relationship;

3. to postulate, on the basis of the explication and theoretically formulated relationship between similarity and learning, a possible resolution of the similarity paradox as expressed in the first question;

4. to test experimentally one of the consequences of the proposed resolution (Hypothesis IV), thereby partly verifying the resolution.

Chapter 3

AN EXPLICATION OF SIMILARITY

The similarity between two stimulus objects (StOs) can be 'measured' in two ways: by analysis or by judgement. In both cases the 'measuring instrument' is the human organism; however, the purpose of the two is different and thus the results are also different. The task of the analyst is to produce a model of similarity while that of the judge is to give an 'estimate' of the similarity.

The task of the analyst is by far the hardest since the usefulness of the model is uppermost on the list of priorities. In considering the similarity between StOs the analyst must decide which are to be rejected. The choice is a necessary one since StOs are overwhelmingly similar to each other. That is, there is always something with respect to which two StOs can be considered as being similar. For example, a shoe and a car are similar in that they are physical objects possessing mass.

Intrinsic and Environmental Characteristics

The usual line of demarcation is drawn between the characteristics which pertain strictly to the StO and those that relate the StO to the environment. This distinction could have originated with Plato but it was certainly used by Hume in his analysis of the similarity between ideas.

To begin with, Hume divided <u>all</u> the "perceptions of the mind"

into two classes: those which were "less forcible and lively" he called "Thoughts or Ideas"; and those which were "more lively" such as when people "...hear, or see, or feel, or love, or hate, or desire, or will" he called "Impressions" (Hume, p. 455). Subsequently, after considering the nature of thought, Hume concluded that all ideas were "...derived either from...outward or inward sentiment..." or philosophically speaking "...all...ideas...are copies of...impressions..." (Hume, p. 456).

Complex ideas being composed of simple ideas were similar to the degree to which they had these in common (Hume, pp. 456, 471, 477; Wallach on Hume, 1958, p. 103). Simple ideas, now reduced to simple sensory impressions, on the other hand, were not aggregates but basic building blocks. How could they then be similar? For example, how could blue be more similar to green than to red? To capitulate by saying that "...similarity between simple ideas is somehow given... directly..." is not a satisfactory solution (Wallach, 1958, p. 103).

The cause of the difficulty, according to Wallach, is that the admissible stimulus characteristics are restricted to those which pertain only to the StO (intrinsic properties). Wallach suggested that the problem could be partly solved by broadening the range of admissible stimulus characteristics, or 'attributes' as he calls them, to include also those which relate the StO to the environment. For example, a table could be defined in terms of "...use (the things done upon it), of location (the place where it is found) or of construction (how it is built): (Wallach, 1958, p. 103).

At first, this seems like a workable solution; however, on closer inspection it makes similarity indeterminable. If the two

types of characteristics are examined one can see that intrinsic characteristics can be determined even for complex StOs by analysis. Furthermore, there is a high probability of obtaining agreement between analysts. On the other hand, the characteristics which relate the StO to the environment often depend on the analyst's past experience and ability to shift frames of reference. Consequently, great diversity of characteristics and very little agreement may be expected from the analysts.

In order to avoid introducing such highly personal variables it seems advisable to retain Hume's restriction that similarity between StOs be analyzed _only_ in terms of the characteristics displayed by the StOs themselves. This is a necessary and useful boundary condition since the object is a model of similarity which is to serve as a reference or basis and therefore must be stable and not variable. However, if Wallach's solution is rejected, how _can_ the similarity between simple ideas be explained?

Common Elements Versus Dimensions

Basically Hume's conception of similarity is finite in character, defining simple ideas as the basic units of comparison. Consequently, if these units (or attributes) seem similar to other units there is no way to define this new similarity. In other words, whereas StOs were similar to the extent that they exhibited a number of common attributes, the similarity between attributes could not be reduced to further attributes. Thus two StOs could be regarded as similar because they both had the attribute 'blue' in common or the attribute 'green' or the attribute 'luminous'; however, what do the attributes 'blue' and 'green' have in common?

An examination of the nature of attributes suggested that they are like concepts in that one class label is used to refer to a set or range of related characteristics. For example, such attributes as 'roundness' or 'luminosity', mentioned by William James, cover a wide range of varying degrees of each characteristic. There is a marked difference between the luminosity of the moon and that of a gas-jet; also between the roundness of a foot-ball and that of the moon. Nevertheless, these differences are subsumed under one label and the StOs are said to exhibit a common attribute regardless of differences in degree.

The great variability in degrees of attributes suggests that they may be regarded as class intervals on a continuum, each interval enclosing a certain amount of some underlying dimension.[1] For example, the attribute blue, defined in terms of wave length, may be thought of as a class interval on the colour continuum, having upper and lower wave length limits of 5.00×10^{-5}cm. and 4.50×10^{-5}cm. respectively.[2] In the same way attributes such as 'hard' or 'luminous' may be regard-

---

[1]The terms 'continuum' and 'dimension' are used to convey two slightly different meanings. A 'dimension' is defined as a measurable extent, and a 'continuum' is defined as an extent, either measurable or not. For example, 'weight' may be regarded as a continuum; on the other hand, the weight continuum in terms of pounds, grams, etc. may be regarded as the dimension of weight. Consequently, if the term 'dimension' is used to refer to a continuum for which a scale of measurement does not exist, then this is to imply a belief that a scale _can_ be developed.

[2]This refers to spectral colours only. All others may be described in terms of three parameters: dominant wave length, reflectance, and purity. Thus the colour continuum may be defined in terms of a measurable extent which is either unidimensional or multidimensional. In other words, a given dimension may be a 'resultant dimension' of two or more component dimensions, or a single dimension to begin with.

ed as class intervals on the underlying dimensions of the 'hard-soft' and 'dull-luminous' continua respectively. Although each attribute spans most of its respective continuum, a 'class-interval' approach is useful since it implies that attributes may be regarded as possessing sub-attributes. For example, the 'soft-hard' continuum may be divided into 'very soft', 'soft', 'semi-hard', 'hard', and 'very-hard'. Although there is no single label in the language either for the individual classes or for the whole 'soft-hard' continuum, this does not deny the existence of the underlying dimension or the possible class intervals on it. It seems that only the colour continuum has been subdivided in the language into many class intervals. However, the colour names do not refer to equal intervals on the colour dimension (wave length). Thus for example, there is a single word 'orange' for the colour 'yellow-red' but no comparable class label for the colour 'blue-green'. The wave length interval, however, is almost identical in both cases. It can be surmised, therefore, that colour names were derived intuitively, depending to a large extent on the local needs for discrimination. For example, in some areas, languages have evolved which do not distinguish between blue and green but have a common label for both (Brown, 1956, p. 303-310).

Consequently, there is nothing rigid about the class labels, and for the purpose of a similarity model the class intervals may be set arbitrarily and provided with new names. This is quite consistent with the usual explication of scientific terms which invariably have different meanings from every day usage (cf. the concept of 'fish').

In summary, similarity between StOs may be regarded as being of the 'common elements' type, and the degree of similarity is determined by counting the number of identical elements, attributes, or

simple ideas. However, if the attributes themselves are regarded as StOs to be compared, then they have to be regarded as dimensions (or contiguous class intervals on the underlying dimensions). In other words, when dealing with attributes it is only necessary to stop considering them as discrete and unrelated entities, and to go back to the basic dimension on which the attribute happens to be the class interval. Thus similarity may be determined either as a proximity on the dimension, or new and smaller class intervals may be formed from the previous large interval and similarity may be determined by the number of sub-classes separating the two attributes in question. Consequently, blue and green are more similar than green and red, because the members of the former pair are approximately 0.5000×10^{-5}cm. apart in terms of wave length, and the members of the latter pair are approximately 1.100×10^{-5}cm. different in wave length. Alternately, blue and green may be regarded as more similar than green and red since the former are contiguous class intervals on the colour dimension while the latter are separated by two intervening class intervals. Although the colour intervals are not equal the example illustrates the procedure which may be used.

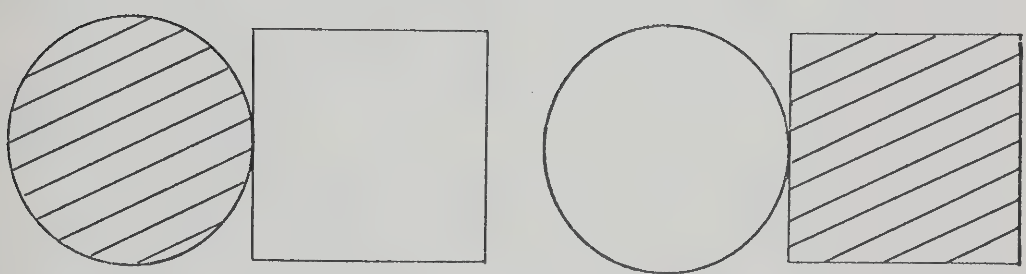
Consequently, the division between intrinsic and environmental characteristics can be maintained and a model of similarity developed strictly in terms of the intrinsic characteristics. Furthermore, within such an interpretation the two major theoretical similarity models ('common elements' and 'dimensional') may be seen to be two aspects of the same approach. Just as in statistics one may work with grouped or ungrouped data, so with respect to similarity the needs and circumstances might be taken as the determining factor in deciding which

approach is better under a given set of conditions. This is not at all contrary to scientific precedent since in physics, for example, an electron may be regarded either as a particle or as a wave, depending on which description is more useful for a given purpose.

A dual similarity model has one other advantage. The major problem associated with the purely dimensional model is that it cannot deal adequately with StOs possessing mutually exclusive dimensions. On the other hand, no such problem is associated with the common elements model. Therefore, a dual model provides a solution to this dilemma; that is, when mutually exclusive dimensions are encountered it is only necessary to regard the dimensions (or parts of them) as discrete entities and to determine similarity by counting the number of dimensions in common.

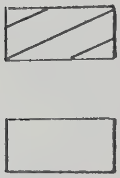
Spatial Relationships

There is, however, one weakness common to both models. Complex StOs cannot always be described in terms of attributes or dimensions that pertain to the whole StO. That is, a given attribute may refer either to the whole StO or only to a part of it. In the case of the attributes cited in William James' famous illustration the former is true, that is, the whole moon is round not just a part of it. This, however, is not always the case. Complex StOs may have parts that differ from each other with respect to colour, geometrical shape, texture, luster, hardness, etc. The example on the following page illustrates this. StOs A and B both exhibit the attributes 'roundness', 'squareness', 'yellowness', and 'greenness'. However, since the attributes refer to different parts of the StOs, they form spatial relation-



StO A　　　　StO B

Key:



- yellow

- green

ships. It is important to consider these since the model by implication defines identity, and if spatial relations are disregarded, the model will lead to inconsistent results. That is, if similarity is determined by a counting of common (that is, identical) attributes or elements, then different degrees of similarity are possible depending on the proportion of common attributes to the total number of attributes. In other words, since StO A and B display four common attributesout of a total of four possible attributes, they should be identical. This, however, is obviously not the case. Consequently, the spatial relationship among the attributes has to be considered as an attribute itself.

This consideration is particularly important since the formal similarity definition almost universally used in verbal learning is a direct consequence of such a model and therefore, has the same problem. The solution is identical.

The Concepts of Identity, Similarity and Difference

No treatment of similarity would be adequate if the related concepts of identity and difference were not also treated.

It is obvious that a StO cannot be similar by itself, it has to be similar with respect to something. Consequently, similarity is a relationship between two StOs and may be considered a 'relation'. Furthermore, a given StO may be more similar to some and less similar to other StOs. Thus various degrees of similarity are possible. If the degrees are ordered according to magnitude then the question of limits arises immediately. What are the end points, if any, of this sequence of degrees? The obvious choice seems to be to consider a

degree of '100% similarity' as 'identity' and hence one of the end points.

However, an argument presented by Noble (1957) based on an analysis of logical structure shows that one cannot consider 'identity' as '100% resemblance'. He defines similarity using 'symbolic language' as "...a relation between two elements' (p. 25).[3] This relation exhibits the properties of nontransitivity, symmetry, irreflexivity, and nonconnexity while the 'identity' relation is transitive, symmetrical, reflexive, and nonconnexive.[4] The lack of isomorphism between the relations on two of the properties is sufficient ground for rejecting the statement that 'identity' is '100% similarity'. Nevertheless, a deeper analysis of Noble's argument is required.

In one of his arguments Noble illustrates a joint by indicating that relations such as 'greater than' and 'less than' are transitive. This is true, as long as they are assumed to connect unidimensional entities. The 'greater than' relation is no longer transitive and in fact has no meaning when applied to such StOs as ordered pairs of those from the set A=(2 lbs, 2 ft, 3 sec, 5 cm/sec). What meaning is there in (6,2)>(2,5) or in 2 lbs > 2 ft? Is the 'greater than' relation transitive or nontransitive in such cases? A relation, it seems,

---

[3]By element' Noble means 'StO' and the latter term will be used henceforth. The term 'element' will be used only to mean 'component part' as in the phrase 'common element'.

[4]The equality relation may be used to illustrate the meaning of the various properties:

Transitivity:	if a = b and b = c, then a = c
Symmetry:	if a = b then b = a
Reflexivity:	a = a is meaningful
Nonconnexity:	a = b is sometimes meaningful, depending on the universe.

cannot be tested for transitivity symmetry or any other property *in vacuo*, that is, without specifying the nature of the StOs.

In his article Noble (1957) does not specify the nature of the StOs but the implication is that a most general approach is intended. Thus it might be assumed that the StOs are members of the real-and-abstract universe. If the dimensional approach to similarity is first considered, then the StOs are either unidimensional or more probably multidimensional. In the case of unidimensional StOs the similarity relation is at all times transitive. This is obvious since all StOs have some degree of a common dimension, and as long as they possess even an infinitesimal amount of that dimension every StO will be similar to every other StO. A practical example is a set of pure tones. Should the StOs be multidimensional, as they usually are outside the laboratory environment, then some rule may be given according to which the various dimensions may be combined into one resultant or index, thus *retaining* the transitive nature of the similarity relation. Nevertheless, one exception does exist.

It was assumed above that all StOs possess, with varying intensities, a common set of dimensions. However, if this is not so and some StOs have one or more mutually exclusive dimensions, then the similarity relation need not be transitive. This is parallel to the 'common elements' approach where StOs have mutually exclusive 'elements' or as in James' example, mutually exclusive 'attributes'. It is quite possible that he was the first to point this out in his classical example of the moon, gas-jet, and foot-ball comparison.

> Moon and gas-jet are similar in respect of luminosity, and nothing else; moon and foot-ball in respect of rotundity, and nothing else. Foot-ball and gas-jet are in no respect similar-

> that is, they possess no common point, no identical attribute (James, p. 378).

Since attributes, at times, may be conveniently replaced by dimensions or parts of dimensions, this example may be taken as illustrating a case of mutual exclusiveness in terms of either attributes or dimensions. Consequently, the similarity relation can be either transitive or nontransitive, depending on the dimensions or attributes possessed by the StOs. In summary, as long as all StOs possess at least one element (attribute, dimension) in common with every other StO in the universe of StOs, the transitivity property holds. This, however, is a most unsatisfactory limitation.

The other property besides transitivity which is not shared by the 'similarity' and 'identity' relations is 'reflexivity'. According to Noble, the statement 'A is similar to A' is never meaningful. However, whether a StO is similar to itself or not is more a matter of definition of philosophical position rather than of logical proof. While Noble chooses to consider the similarity relation irreflexive, others may choose to define it as reflexive. Nevertheless, this question is an academic one since the similarity relation can only be transitive under severe limitations and therefore, there is no need to try to define it as reflexive.

Consequently, under the most general circumstances, in both models of similarity, the 'identity' and 'similarity' relations are not isomorphic and there seems to be

> ...no justification for interchanging them within any formal context. To do so would constitute equivocation and result in chaotic discourse. Those who consider identity as '100 percent resemblance' may derive this from some Platonic agrument about ideal limits, but not on the basis of an analysis of logical structure (Noble, 1957, p. 27).

Nevertheless, the acceptance of the two relations as not isomorphic and mutually exclusive does not ipso facto invalidate the statement that '100% resemblance' is 'identity'. The similarity relation as formulated by Noble has the form 'similar'; the above relation, on the other hand, has the form 'n% similar'. The two relations are not the same: one is qualitative and the other is quantitative. The former treats similarity on an all or none basis, that is, two StOs are either 'similar' or they are 'not-similar'; the latter, differentiates between varying degrees of similatity within the connective. Therefore, 'n% similar' may be interpreted as a compound relation joining two simple and mutually exclusive relations with the 'exclusive or' connective. Thus what 'n% similar' stands for is:

$$[\text{n\% similar} \mid 0 < n < 100 \quad \text{(OR)} \quad \text{identical} \mid n=100]$$

In other words, the StOs are either similar ($0 < n < 100$) but not identical ($n=100$); or they are identical ($n=100$) but not similar ($0<n<100$). In this way 'similarity' and 'identity' remain mutually exclusive and since there is no requirement of isomorphism for the formation of a compound relation, Noble's conditions are met. Consequently, there is no question of interchanging 'similarity' with 'identity', but a matter of combining the two in an 'exclusive or' relationship. The statement '100% similarity is identity' although technically inaccurate is, nevertheless, connotatively expedient. Perhaps the precise expression should be used:

$$[\text{n\% similar} \mid 0 < n < 100 \quad \text{(OR)} \quad \text{identical} \mid n=100]$$

However, as long as 'n% similar' is regarded as being thus defined, there seems little harm in using the shorthand notation. There is, in fact, an added advantage since the precise expression has to be modified even further if it is to include the second end point of 'difference' when 'n' equals zero percent. The precise expression would be too cumbersome since it would have to express the following logical relationship:

$$
\begin{array}{lll}
(A \oplus B)C' + A'B'C & \text{where } A=\text{similar} & (0<n<100) \\
 & B=\text{identical} & (n=100) \\
 & C=\text{different} & (n=0) \\
 & A', B', C' \text{ indicate negation of } A, B, \& \ C & \\
 & \oplus \ \text{'exclusive or'} & \\
 & + \ \text{'or'.} &
\end{array}
$$

Since this is a relatively complex statement it is much simpler to use 'n% similar' and to define it to mean the above. Thus there is no question of interchanging similarity and identity but only an attempt to use a shorthand notation.

The Concept of 'Level of Detail'

The relationship between similarity and difference has been clarified and reduced to an exclusive 'or' relation. Nevertheless, there is one qualification that has to be made.

Usually when the term 'different' is used outside the laboratory environment there is little disagreement as to the relative correctness of the designation, since most individuals assume a 'common level of detail'. However, there _are_ different levels, and StOs that are different at one level may be either similar or identical at a different level of detail. As early as 360 B.C. Plato pointed this

out, everything

> ...when seen indistinctly and at a distance, appears to be one; but when seen near and with keen intellect, every single thing appears to be infinite...as in a picture things appear to be all one to a person standing at a distance, and to be in the same state and alike...[but when approached]...they appear to be many and different (p. 511).

This implies that StOs may be perceived at different 'distances' or levels, and those that were considered identical at one level (of detail) would be different at a finer level. For example, an 'A' is still an 'A' whether large, small, printed, written, or painted. However, at a finer level of detail the exact physical appearance of the letter may make it different, or only similar to another 'A'. Consequently, questions about similarity, difference, and identity have to be modified to include a reference to the level of detail at which a given analysis or judgement is to be made. This is particularly important for learning experiments where non-meaningful materials are used since the <u>Ss</u> have to work at a finer level of detail than is usual when the StOs are meaningful. In other words, meaningfulness may itself be regarded as being on a different level. For example, evaluation of similarity between meaningful words does not involve the examination of the individual letters for identity. At the other extreme, evaluation of similarity between two letters may be reduced to considering either their visual or aural characteristics. Thus, the level of detail is not a linear hierarachical representation but rather may involve the use of more than one modality at essentially the same level.

Consequently, the similarity relation is not absolute but rather <u>relative</u>, and the question, whether two StOs are similar <u>cannot</u> be answered without asking <u>at what level of detail</u> the judgement or

analysis is to take place.

The Relationship Between the Theoretical and Empirical Models

As was stated at the beginning of the Chapter, the task of the analyst and judge are rather different, and so are the products of their work. The theoretical model can use, in fact must use, intrinsic characteristics of the StOs in arriving at evaluations of similarity. However, no such restriction is placed on a judge who is often asked to estimate the 'overall similarity'. Furthermore, the analyst must specify at what level of detail the evaluation is being made; a judge, on the other hand may use any level that seems appropriate to a particular comparison. Thus, it is quite possible that different pairs of StOs are judged at different levels.

The relationship between these two ways of 'measuring' similarity is particularly significant for learning studies where the definition of similarity that is presently being used is strictly in terms of the physical or intrinsic characteristics of the StOs. The Ss on the other hand, may behave more like the judges, that is, they may react to the environmental characteristics rather than to the physical ones. Furthermore, the Ss may at times use a different level of detail from that which E expects them to use. For example, the researcher may consider that in the set (BLK, WHT, WHL) the first two trigrams are different while the last two are similar. The Ss, on the other hand, may perceive the first two as also being highly similar because they both contain the essential letters of the polar pair BLacK and WHiTe. Thus, while for the experimenter one pair is similar, for the Ss two pairs of trigrams may be similar. Consequently, the degree

of similarity is much higher than would be expected strictly in terms of physical similarity! Nevertheless, the experimenter has no other choice but to use physical similarity because psychological similarity can be expected to be too unstable a variable. This is evident from the fluctuations to be found in similarity judgements which vary not only from S to S or from one group of Ss to another but are also influenced by such factors as the relative position of the StOs during the presentation or by the nature of the whole stimulus set. For example, Attneave (1950) found that the similarity judgements were affected by the physical distance between the StOs on the cards on which they were presented.

> More specifically stated, the perceived similarity of two triangles depends not only upon similarity of slope between corresponding sides, but also upon similarity of slope between proximate sides. Thus the perceived similarity of the same two triangles differing in form may either be enhanced or decreased by a change in their left-right placement (p. 542).

The difficulty with similarity judgements is that they 'do not stay put' as Torgerson (1965) would say; that is, "...similarity does not exist as a unique, invariant relation between a pair of stimuli, but rather depends upon such things as stimulus context and the cognitive strategy taken by the subjects" (p. 383). Thus in one experiment described by Torgerson (1965) some of the Ss used two psychological dimensions in arriving at their similarity judgements while the rest used three. The different strategies, of course, resulted in 'very different configurations' (p. 386). As Wallach (1958) pointed out in his article on similarity, Ss may perform a 'selecting and ignoring function' and thus focus on different dimensions or combinations of dimensions in judging similarity. Furthermore, it cannot be assumed that all the

dimensions will be given equal weight by all the Ss, nor that all Ss will be even aware of all dimensions in the first place. Chambliss (1957), for example, found that physical measures on the sides of geometrical figures were not significant while those on the angles were very important in determining similarity judgements. Although in the latter case there was consensus among the Ss as to which characteristics of the StOs were to be weighted more, this is not always the case (cf. Torgerson, 1965). Furthermore, in 1962, Osgood found that even though the Ss were told the three dimensions that they were to use in their judgements, three different groups of Ss produced different results depending on their cultural background. But, as Muir (1968) has pointed out, these results have mostly been ignored by researchers including Osgood. Nevertheless, the fact remains that different Ss judge similarity differently, and the same StOs may be judged to have different degrees of similarity under different conditions. If this is the case when the attention of the Ss is specifically directed to similarity, how much greater must the variability be when the Ss are simply learning the material. This is, perhaps, one of the reasons why experimenters have attempted to study only the relationship between physical similarity and learning. However, if it is psychological similarity to which the Ss are responding, should it not be used as the independent variable in spite of the variability? Perhaps, however, the idea does not seem to be too practical since it would have the tendency to anchor research in a morass of variability. It seems, therefore, much more meaningful to keep on striving to produce better and better definitions and models of physical similarity and to use this reference system for trying to understand how different

Ss see different things as being similar and how different circumstances produce different perceptions of similarity. Such a reference system would 'stay put', and would be a unique and invariant relation between a pair of StOs.

Consequently, one way of looking at the relationship between physical and psychological similarity is to consider the latter as an estimate of the former. That is, given the task of judging StOs for similarity, what the Ss produce is only an estimate of physical similarity because they are required to give an 'overall' evaluation without benefits of any measuring instrument other than themselves. Consequently their estimates, the psychological similarity judgements, are understandably rather inaccurate. Like any other judgements they are subject to any and all perceptual errors and are highly dependent on the Ss ability. Osgood (1962) showed that cultural background was a factor in similarity judgements and therefore, it is very likely that there are other variables such as age, sex, intelligence, etc. that might also affect the quality of the judgements. In psychophysics it is acceptable to talk about the accuracy of judgements when the relation is heaviness, for example. Why should it be less reasonable to do so for the similarity relation? In both cases the Ss are comparing a set of StOs. Perhaps the reason is the lack of a measuring device for similarity. When a S judges weight A to be heavier than weight B this can be checked immediately for accuracy because a measuring device exists which will, within a slim margin of error, give a reasonably accurate answer. On the other hand, there is no ready made instrument which could be used to 'measure' the degree of similarity between two StOs, and the best approximation can only be

obtained by painstaking analysis. The theoretical approach to similarity, of course, is trying to set up some sort of guidelines whereby similarity between StOs could be determined easily and with as small a margin of error as possible. However, the task of separating reality from illusion and any effects caused simply by variability in the Ss' ability or the external conditions is not easy. Nevertheless, at one time it was thought that the sun went around the earth and although the perceptual illusion still persists, this is no longer believed.

This then is one way of looking at the relationship between physical and psychological similarity.

Chapter 4

THE RELATIONSHIP BETWEEN SIMILARITY AND LEARNING

The application of the explicated concept of similarity to learning required an examination of the other variables relevant to the learning situation. First, it was necessary to determine the basic learning behavior of the subjects (Ss); that is, are the Ss primarily passive or active participants in a learning experience? An examination of the relevant literature indicated overwhelming evidence in support of the hypothesis that Ss actively look for learning strategies and for means to facilitate the learning task, and that they do so under all sorts of learning conditions. Therefore, if the Ss are active seekers of learning strategies, what would they do with similarity? An examination of similarity in the context of learning indicated that similarity has a tremendous potential for serving as a basis for organization. This raised the next question, what is the effect of organization on learning? The evidence from research supported the hypothesis that organization facilitates learning. Consequently, if Ss are active learning-strategy seekers, and if similarity has a great potential for organization (which is known to be a facilitative factor in learning), then the conclusion that similarity should facilitate learning is inescapable. However, this conclusion was contrary to at least half of

the findings of verbal learning experiments.

This chapter will summarize the literature dealing with the three crucial hypotheses (propositions) which led to the conclusion that similarity should facilitate learning. The next chapter will examine this conclusion in the light of the experimental findings thereby explicating the similarity concept in relation to learning.

Assumption. Similarity is one phenomenon. Therefore, under equivalent conditions it should affect learning in the same way.

Proposition I. Ss do not primarily learn by rote. Given a learning task, whether paired-associate (P-A), serial, or free recall, Ss seek devices or strategies whereby they can make the task easier. They tend to organize, structure, or make sense of the material so as to facilitate learning. They do so even under the most adverse conditions.

In other words, the Ss are not passive but active participants in the learning task set before them.

Proposition II. Similarity, being a relationship between two or more stimulus objects (StOs), has potential for organization.

A given StO cannot be similar by itself but must be similar with respect to something. This relationship, therefore, may be used as a basis for grouping the StOs into classes. Since there is no guarantee that Ss will use the common attributes, dimensions, or relationships for classification; therefore, it must be said that similarity provides only an opportunity for organization or, in other words, has a potential for organization.

Proposition III. Organization of the material to be learned facilitates learning.

The term 'organization' is to be understood to refer to such activities as grouping, 'chunking', coding, clustering, structuring, ordering alphabetically, using mediators or mnemonic devices, spatial configurations or patterns, etc.[1]

Conclusion. Similarity should facilitate learning assuming that it is one phenomenon.

More specifically, similarity, assuming that it is one phenomenon, whether among meaningful or nonsense StOs, will facilitate learning *if* the *Ss* are given the opportunity to use the organizational potential due to similarity in the material.

Proposition I

> *Ss* do not primarily learn by rote. Given a learning task... *Ss* seek devices or strategies whereby they can make the task easier (p. 56).

Similar propositions have been put forth by Gibson (1966), Manis (1966), and Shepard (1966). Their articles represent a revival of interest in an area which has been ignored since the time of Gestalt Psychology.

> It has been ignored because the traditions of, first, behaviorism and, later, operationism tended to prevent psychologists from speculating about symbolic processes inside the memorizer and encouraged an organization of the field in terms of what the experimenter, rather than

---

[1]For a discussion of 'chunking', coding, clustering, etc., see pp. 73-79.

> the memorizer, was doing....[however]...It is the learner, not the experimenter, who must be explained (Miller, Galanter, and Pribram, 1960, pp. 126-126).

The authors go on to a severe, and at times, sarcastic criticism of experimental psychologists who prefer to ignore the existence of a mediating or organizing process in the learner. In this they find themselves in agreement with the criticism of Woodworth and Schlosberg expressed as early as 1954.

Since that time, but more specifically since the early sixties, evidence has been accumulating in support of the above proposition. Therefore, it seems that the learner can no longer be viewed as a passive system, but as

> ...an active agent with a definite, hopefully decipherable, internal structure. The ubiquity of evidence for processes of searching, grouping, and ordering in recent studies suggests...that these processes represent neither occasional lapses nor mere epiphenomena of this underlying structure. They represent...its fundamental modus operandi (Shepard, 1966).

In P-A learning, for example, there is little doubt that mediation does occur. "Subjects too consistently report the use of verbal mediators to link the stimulus representation and the response representation to deny the basic process" (Underwood, 1964). Reports of such strategies can be found as early as 1947 (Heidbreder, pp. 112-113) and as recently as the 1960's (Underwood and Schulz, 1960; Farber, 1963; Eagle and Leiter, 1964; Martin, Cox, and Boersma, 1965; Martin, Boersma, and Cox, 1965; Battig, 1966).

Evidence of grouping processes comes from free recall studies where a phenomenon called 'clustering' was observed. Bousfield and Bousfield (1966) define this as "...The occurrence of sequences of related items in the free recall of a randomly ordered stimulus

list" (p. 935). The lists usually contain words from a number of categories such as animals, vegetables, professions, articles of clothing, or musical instruments, etc. In recall it is found that the Ss group like items together and that this grouping occurs significantly more often than it should by chance alone (Bousfield, 1953; Bousfield and Cohen, 1953; Bousfield, Cohen, and Whitmarsh, 1958; Coffer and Bruce, 1966; etc.). Subsequent research has shown that Ss can use other than the taxonomic classes as bases for clustering (Jenkins and Russell, 1952; Jenkins, Mink, and Russell, 1958; Rothkopt and Coke, 1961; Bousfield and Puff, 1964; Pollio, 1964; Marshall, 1967). A rather comprehensive summary of the different bases for clustering can be found in Bousfield and Bousfield (1966, p. 940).

A different type of clustering was examined by Tulving and his associates at Toronto (see references in Tulving, 1962b, p. 189). Rather than provide Ss with lists containing experimenter (E) imposed categories or associations, Tulving used randomly selected words, and measured the constancy with which Ss used the same sequential ordering of words in successive free recall protocols. He found that Ss do indeed have a strong tendency to recall items in the same order on different trials even though the words are presented in a different order on every trial (Tulving, 1962a). Since it is the S who imposes this ordering, Tulving called it 'subjective organization' or (SO). A similar procedure was later adopted by Bousfield and his associates; however, they called it 'intertrial repetition' or ITR. Their findings confirmed those of Tulving and his associates (Bousfield, Puff, and Cowan, 1964; Bousfield and Abramczyk and Bousfield, 1967).

Glanzer and Meinzer (1967) attempted to determine the nature of the organizing process during a free learning task. Their conclusion was somewhat negative in that they found that simple repetition of the words to be learned was not an effective technique and lowered overall recall; therefore, an effective technique must "...consist of some other activity" (p. 934). However, they do not speculate as to its nature.

Kendler (1964), discussing concept formation, touched on the topic of clustering since he felt that its implications were relevant to conceptual behavior. In fact he concluded that "...the phenomenon of clustering illustrates concepts in action" (p. 221). Miller (1958) made an even more general statement when he said that "...verbal learning is merely a simple form of concept formation" (p. 490). Osgood would probably take exception to these statements and consider such 'concepts' much too broad to have any useful function (see Osgood, 1953, p. 667). In his criticism of Hull's classical experiment (1920), Osgood objected to the use of the term 'concept formation', stating that it would seem that Hull was "...studying the development of <u>labeling</u>...." rather than that of concept formation (p. 667). However, Bruner, Goodnow and Austin (1956, pp. 41-43) list three types of concepts: conjunctive, disjunctive, and relational; and it is the formation of conjunctive concepts that Hull was studying. Thus, the arguments revolve around the selection of an appropriate name for this type of behavior, not as to its nature. However, regardless of how it is named, this type of activity still remains a form of structuring or organization and manifests itself as clustering in free recall studies. In Hull's experiment <u>Ss</u> must also

have searched for means to organize the material in what they were told was a straight memorization task since, with practice, they were able to correctly label new instances of the concept.

More recent evidence that Ss do not as a rule learn by rote comes from studies by Segal and Halwes (1965, 1966), and Smith (1965). Segal and Halwes (1966) found that when Ss were presented with a relatively large number of verbal sequences the Ss made 'errors of comission'. That is, they generated many sequences which were not in the original list but which could be produced on the basis of a simple 'generative grammar' (cf. Chomsky and Miller, 1963). The authors conclude: "In addition to learning some of the sequences the Ss learn rules of combination by which they generate sequences" (Segal and Halwes, 1966, p. 157).

Further evidence of Ss' perspicacity comes from a host of studies on stimulus selection (Shepard, Hovland, and Jenkins, 1961; Sundland and Wickens, 1962; Underwood, Ham, and Ekstrand, 1962; Lawrence, 1963; Underwood, 1963, pp. 33-48; Leicht and Kausler, 1965; Martin,1967). As the recency of these dates indicates, stimulus selection (or the difference between the nominal and functional stimulus) was not considered relevant to the standard S-R model prior to the 1960's. However, by the time of the second conference on Verbal Learning and Verbal Behavior its status had changed. Wickens (1963), summarizing the conference stated:

> ...the concept of the functional stimulus entered into the previous conference only in a tentative and speculative fashion. This time it has been represented by a formal paper, and this quasiperceptual process - stimulus selection by the subjects - seems to be well integrated into the thinking of S-R theorists....the term was used quite freely and in an explanatory fashion....It seems...that there was a bold, confident usage of a concept that would have been used appologetically-

> or beligerently in the form of an abstraction - in circles such as this only a few years back (p. 377).

It was only in 1958 that the full importance of the implication of discrepencies between the 'nominal' and 'functional' stimulus came to Underwood's attention (Underwood, 1963, pp. 34-35). The observations were made by Mattocks under Underwood's direction and are reported in Underwood and Schulz (1960, 297-298). From post-experimental interviews Mattocks discovered that 62% of the Ss reporting associations believed that they had used only the first letter of the trigram as the functional stimulus.

> ...it is evident that in many learning situations, upon presentation of S [nominal stimulus], Ss selectively focus on only certain aspects of S [functional stimulus] and that frequently they code or recode those perceived aspects in various ways so as to obtain a facilitory mediating link between S and R....(Martin, 1967, p.272).

In fact Underwood (1963) commenting upon this selective behavior on the part of the S states that

> ...The college sophmore is a perfect confirmation of the law of least effort. Why fuss around with all three letters of the stimulus when a single one will do the trick? (p. 35).

The fact that stimulus selection may occur is definite evidence that Ss cannot be considered as passive organisms but, on the contrary, must be regarded as active participants who bring to the situation well-developed learning habits, preferences, and strategies. That Ss prefer to deal with certain stimuli over others has long been recognized in rote learning studies (Korwoski, 1931; Postman and Riley, 1957), and in studies on concept formation (Eckstrand and Wickens, 1954; Shepard, Hovland, and Jenkins, 1961). As early as 1939 Newman demonstrated that not all parts of a prose passage were equally important to the learner. Subsequently,

Postman and Senders (1946) showed that what is selected as important and what is remembered is strongly determined by the Ss' set, and that the latter can be manipulated to produce a variety of selections of information. In other words, when Ss are presented with a situation they will

> ...usually interpret the situation as requiring a particular 'mode of attack' and...these 'demand' characteristics of the situation, without very specific instructions, determine the way the subject will handle the task (Mandler, 1967, p.19).

In the absence of instructions Ss will make up their own. For example, in one free recall task, without specifically being told to do so, two Ss reproduced the word list in alphabetical order (Tulving, 1962b). There is, of course, no guarantee that any two Ss will interpret the situation in the same way, select the same mode of attack, or use the same technique of structuring the material. Discussing the related topic of Ss' awareness of the purpose of the experiment, Feber (1963) suggested that Ss should be questioned on this point post-experimentally, and that this 'awareness' should be considered as a variable in many instances. It seems reasonable to conclude that a similar procedure should be used to find out what the Ss' learning strategy was, particularly in view of the evidence that Ss as a rule do not seem to learn by rote but look for relationships, structure, and means to facilitate even the most meaningless learning tasks. At times, it seems that Ss find regularities or rules even 'when none are to be found'. Shepard (1963, p. 61) cites a study by H. M. Jenkins (reported at the 1960 meeting of the AAAS) in which Ss were presented with pairs of digits and instructed to produce a third digit as a response. However, since the responses were said to be correct or incorrect according to a prearranged schedule, no

'solution' was possible. Nevertheless, the _Ss_ almost invariably insisted that they had discovered rules, although typically complex ones, which would enable them to give the correct responses.

Proposition II

> Similarity, being a relationship between two or more StOs, has potential for organization (p. 56).

This proposition may seem somewhat self-evident. However, what may not be immediately clear is that almost all of the evidence cited under Proposition I illustrates grouping or organization on the basis of similarity.

Clustering on the basis of common categories as well as on the basis of synonymity is rather an obvious example. However, associative clustering is somewhat less obviously similarity based. Osgood (1953), discussing this problem, states:

> All semantically determined associates are _similar_ in some way to the stimulus word, similar in _meaning_...similar in _context_...or similar in the sense of _hierarchical relations_ (NEEDLE _steel_, _instrument_, _implement_, _tool_, _metal_)(pp.708-709).

So called 'clang' associations although not meaningfully similar are accoustically similar. Karwoski and Schacter (1948) conducting an extensive analysis of free associates found that these could be categorized as some form of either similarity or contrast. Osgood (1953) attributes the production of the latter to verbal habits in the culture and to overlearning. Nevertheless, there is an alternate interpretation. Contrast associations such as light-dark, good-bad, or man-woman may be regarded as being highly _similar_ in that they are, in a sense, symmetrical representations or mirror images of the same thing. That is, they are identical except in that they occu-

py opposite ends of the semantic spectrum or are rotated 180° on the dimension of meaning. Such an interpretation would be consistent with the high similarity ratings that Ss seem to give to mirror-image nonsense words and geometrical figures (Runquist, 1968a & b; Attneave, 1950). In addition, Haagen (1949) found that ratings of similarity and ratings of association are highly correlated (pp. 290-291). Consequently, all associative clustering, and therefore all clustering, may occur as a result of similarity based grouping.

In P-A tasks the grouping is superimposed by E and Ss have to learn which two words belong in the same group (pair). An examination of mediational strategies used by Ss in one study by Martin, Boersma, and Cox (1965) showed that 59% of these were similarity based. That is, 24% of the Ss used the common part of each word to remember the pair and 35% of the Ss invented an intervening word which was similar in some way to the two that were to be associated. The remaining 41% of the Ss either used no strategy that they were able to verbalize (12%), used simple repetition (11%), or invented a complex coding strategy (18%). Thus it seems that more than half of the Ss found similarity useful in remembering the P-A pairs. When the similarity of a pair is considered against the ground of the other pairs in the list it seems that Ss prefer to learn like or similar pairs first even when this similarity is only group membership. Thus in a study by Postman and Riley (1957), Ss tended to learn number-number pairs and word-word pairs first and mixed pairs (number-word) second.

However, the best example of the organizational potential of similarity is to be found in language itself. The human organizm is surrounded by an ever changing environment which is composed of a vast

array of discriminably different objects and people. The Optical Society of America (1953) estimates that there are seven and a half million just noticeable colour differences alone, and in a relatively short time, one comes in contact with a fair proportion of them. Any attempt to develop a vocabulary to match this capacity for discrimination would soon overwhelm the organism. In the words of William James, "The world would become a "buzzing, blooming confusion". However, the organism is saved from chaos by the ability to categorize, that is,

> ...to render discriminably different things equivelent, to group the objects and events and people...into classes, and to respond to them in terms of their class membership rather than their uniqueness (Bruner, Goodnow, and Austin, 1956, p. 1).

Consequently, there are only six colour names in common usage instead of over seven million. From childhood the organism learns innumerable labels or concepts such as 'dog', 'cat', 'friend', 'house', etc.[2] In each instance the label refers to a wide variety of discriminably different StOs which, nevertheless, are grouped together because of some common characteristic or property. A child may not be able to verbalize what the critical attribute or set of critical attributes

---

[2]As was mentioned before under Proposition I, some writers refer to this type of learning as concept formation (Bruner, Goodnow, and Austin, 1956; Kausler, 1966; Manis, 1966), others prefer to call it 'a learning of labels' (Osgood, 1953), while still others discuss the controversy without taking sides (Kendler, 1964; Forgus, 1966). Since the controversy is not relevant to the discussion, the author will use the terms 'simple concept' and 'label' synonymously, and the word 'concept' will be understood to refer to such 'labels' as 'dog', 'friend', etc.

is, but as long as it can indicate new instances of the concept it will be assumed to have mastered that concept. Even adults may have to pause and think before they are able to verbalize the _exact_ attributes or dimensions which enable them to recognize members of the very common 'chair' category. To say that these StOs are used for sitting is not enough (e.g. benches, stools, boxes, etc.), nor can they be described as having four legs and a back (e.g. sofa, hanging chairs, one leg chairs, etc.). In fact the task is not as easy as it first appears, and when such classes as 'beautiful things' or 'ugly things' are considered, the problem seems very complex indeed. Hull (1920), patterning his experiment on the natural process of concept attainment, found that although the _Ss_ were able to identify new instances of the conjunctive concepts presented to them, some were unable to verbalize the rule which enabled them to do this.

In addition to conjunctive concepts which are defined "...by the _joint presence_ of the appropriate value of several attributes" (p. 41), Bruner, Goodnow, and Austin (1956) also define disjunctive concepts which, on first inspection, do not seem to be similarity based. A disjunctive category is characterized by a selection rule of the form 'property A or property B or both'. Thus, a 'strike' in baseball is a good example of a disjunctive concept since it is called under three different sets of circumstances. The little research that is presently available suggests that _Ss_ experience great difficulty in learning disjunctive concepts, and often act as if these were conjunctively organized even though they are not given any reasons to assume this (Manis, 1966).

> It is as if people generally expect that when different exemplars are given the same label, they must have some common underlying characteristic(s), and thus they find it difficult to accept the notion that two exemplars may both belong in the same category but for different reasons (Manis, 1966, p. 45).

Bruner, Goodnow, and Austin (1956) cite a number of examples of disjunctive concepts from culturally defined areas, medicine, and science. They remark, however, that in science in particular such concepts are regarded with suspicion, and often are later redefined in conjunctive terms.

> So frequent have been these instances of eventual obsolescence of disjunctive definitions of classes of events that the very discovery of one such seems to act as a challenge to future research rather than as occasion for celebrating a "finding"....One eventually begins to wonder whether Nature herself does not abhor disjunctive groupings! (Bruner, Goodnow, and Austin, 1956, pp. 161-162).

Although, to the author's knowledge, there have not been any attempts to estimate the proportion of disjunctive concepts in the language, it seems that they represent a very small and possibly shrinking percentage of all possible concepts. In view of their relatively small proportion in comparison to the number of conjunctive and relational concepts, they cannot be thought to affect the hypothesis in a significant way. Nevertheless, the fact remains that StOs are placed into disjunctive categories on the basis of a common selection rule, and this may be interpreted as being the one thing that these StOs have in common and thus they are being grouped on the basis of similarity. Although the same argument can be used with respect to relational concepts which are defined "...by a specifiable relationship between defining attributes "(Bruner, Goodnow, and Austin, 1956, p. 43), there is no need to do so.

A relationship such as 'the first digit is larger than the second' used to define the group 71, 43, and 95 can be considered as existing in the StO configuration very much like any other simple physical attribute such as red or round. In fact, in many instances it is difficult to say if a concept is predominantly conjunctive or relational. For example, can a pile of mechanical parts be considered or recognized as a member of the category 'motor', or a stack of prefabricated parts as a 'house'? It seems that "...the joint presence...of several attributes" is not enough to specify such categories as 'dog', 'house', 'motor', 'chair', etc. What is needed in addition to the "presence of attributes" is some specification as to the relationship among the parts or attributes which may be spatial, temporal, or even abstract in character, as in the above case. An example of a temporal relationship is the category or concept 'melody' where a collection of specific musical tones is not enough to define the class. The same is true of the category 'O Canada' and other specific melodies.

As a result of this interpretation, a very interesting question arises: Are all conjunctive concepts or categories relational in character? A quick inspection of the nature of attributes indicates that this is not so. It is found that in those instances where the attributes refer to the whole StO they do not, and in fact cannot, form any relationships with other attributes.[3] In those cases

---

[3]This question will be discussed in greater detail under Proposition III because it is much more relevant there than here.

where the attributes refer only to a part of the StO, relationships are unavoidable and, in the majority of cases, must be regarded in Bruner's terms as "critical or relevant" to the process of classification. That is, without such relationships the StOs would not be recognized as members of that particular class.

In summary then, conjunctive concepts, with or without relational overtones, and relational concepts may be regarded as being similarity based. Disjunctive concepts, on the other hand, may or may not be considered as similarity based depending on the acceptance or rejection of the above given interpretation. However, because of the latter's small proportion, the contention that language is primarily similarity based will not be unduly affected should the interpretation be rejected.

The human organism, surrounded by a vast number of StOs, possessing a great capacity for differentiation, uses the common attributes or relationships in order to organize the enviroment into manageable portions to which distinct labels called 'words' are assigned. However, the vast number of StOs possess many distinguishable characteristics or attributes and if different ones are selected as being 'critical', different groupings may result.

> Obviously, any object may be classed with other objects, depending upon the level of abstraction chosen and the particular set of concepts chosen within that level. For example, a single automobile may be described in a variety of ways, depending upon the level of abstraction chosen and the particular way in which...[one wishes]...to slice that level of abstraction. It may be variously termed a Ford, a convertible, a means of transportation, a blue object, a mechanical contrivance,...a mass of electrons, etc. There are almost an infinite number of ways to describe the car, to class it with other objects...(Forgus, 1966, p. 310).

For each concept there are critical defining attributes and relation-

ships which are shared in common by all members of that class and which distinguish them from non-members. What is a non-critical attribute for one concept may be critical for another. Thus the same StOs may be classified in different ways depending on the choice of critical attributes. The word 'choice' is used advisedly. Bruner, Goodnow, and Austin (1956, p. 7) argue that word categories are not discovered so much as invented.

> Science and common-sense inquiry alike do not discover the way in which events are grouped in the world; they invent ways of grouping. The test of the invention is the predictive benefits that result from the use of invented categories... Do such categories as tomatoes, lions, snobs, atoms, and mammalia exist? In so far as they have been invented and found applicable to instances of nature, they do. They exist as inventions, not as discoveries (Bruner, Goodnow, and Austin, 1956, p. 7).

The student of foreign languages has probably discovered that "... different languages parcel out the same reality in slightly different ways" (Forgus, 1966, p. 311). Thus, for example, such words as 'Zeitgeist', 'savoir-faire', or 'simpatico' lose some of their meaning in translation. Other languages have even greater differences. The Hopi language, for example, groups together under one label all flying things with the exception of birds, while the Eskimo language has three different names for snow (Manis, 1966, p. 87). Kluckhohn and Leighton (1946) found that the Navaho language employs object classes which make distinct grammatical categories for words which denote 'long thin objects', 'round objects', granular substances', etc. The natives of the Torres Straits, on the other hand, fail to distinguish between blue and green in their language. Brown (1956, p. 303-310), discussing this aspect of language and its categories, lists many more examples of "different ways of parcelling out the same

reality". However, all these classes or categories have one thing in common: they are used to refer to StOs possessing common characteristics or, in other words, StOs that are similar to each other. Bruner, Goodnow, and Austin (1956) state that "...there exists a near infinitude of ways of grouping events in terms of discriminable properties..." (p. 7). Kendler (1964) goes so far as to imply that everything is similar to everything else.

> Although it is easy to specify the similarity between two stimulus patterns, it is extremely difficult to decide when two particular patterns fail to possess a feature in common. A coffee can and doughnut possess roundness. What in this world is completely different from either? A shoe? It is definitely not round but it does possess physical characteristics in common with the coffee can and doughnut. They all have weight and substance. They all are _something_. If one argues this way, then only nothing is uniquely different from something. But even here an objection might be raised by some philosophers that nothing is something, or at least the concept of nothing depends on the concept of something (p. 219).

Without going to the philosophical extreme, it is nevertheless possible to appreciate _the vast potential that similarity has for organization_. Not every language will select exactly the same common attributes and relationships at exactly the same level of abstraction or detail to serve as the critical or defining characteristics and relationships for comparable concepts or labels. Nevertheless, the very fact that almost all linguistic categories use similarity for their construction heavily underscores the great importance of similarity as an organizational tool whose applicability is limited only by the _Ss'_ ingenuity in using newly perceived similarities so as to invent new categories and the _Ss'_ facility in using old similarities within the framework of old groupings. Thus, the potential exists. Its utilization, however, depends on the subjects.

Proposition III

> Organization of the material to be learned facilitates learning (p. 57).

In 1940 Katona stated that "...organization is a requirement for successful memorization" (p. 249); more recently Osgood (1953) affirmed: "There would be little disagreement with the statement that organization facilitates learning..." (p. 566); and in 1966 Manis declared:

> ...material that is well structured and tightly organized will generally be recalled far more successfully than that which does not possess a meaningful structure (p. 21).

Evidence supporting this point of view is abundant and comes from many different types of learning experiments. For example, it has been shown that grammatically structured words or nonsense syllables are learned significantly faster than their scrambled counterparts (Miller and Selfridge, 1950; Epstein, 1961; Marks and Miller, 1964; Forster, 1966). The same holds true for letters, that is; the closer the letter groupings approximate English the greater the facilitation effect (McNulty, 1966).

Another common form of organization built into the task by E is blocking; that is, grouping together similar or related words, nonsense syllables, or P-A pairs. This differs from the standard method of presentation in that the related terms are always presented in a contiguous sequence although, within the group, they may be randomized from trial to trial. The results of such studies indicate that blocking facilitates recall (Gagne, 1950; Rotberg and Woolman, 1963; Dallett, 1964; Cofer and Bruce, 1966).

Grouping as a means of facilitation, of course, is not a new

idea. In 1940 Katona compared the recall of Ss who had learned strings of digits subdivided into groups of three with that of Ss who had learned an unbroken string, and found that the former had superior recall (pp. 9-10; pp. 188-189). In 1956 Miller proposed his classical 'chunking hypothesis' and, using binery digits, showed that by recoding a string of ones and zeros into octal notations, Ss could recall about thirty digits or bits as they are often called.[4] According to Miller's hypothesis the human organism can recall 'seven plus or minus two' items or bits. However, if the bits are recoded into 'chunks', of a few bits each and these in turn are recoded into bigger 'chunks', then the organism can indeed recall a very large number of bits since the same 'magic number', 'seven plus or minus two', holds for 'chunks' as well as for bits.

There are, of course, different recoding techniques available ranging from simple grouping as in some of Katona's experiments to coding in terms of concepts. The former, however, is not as effective as the latter since there is nothing in the items to hold them together, at least nothing built in by the experimenter (Katona, 1940, pp. 188-189). When two items, such as the stimulus and response members of a P-A pair, for example, are related either by similarity or compatibility, facilitation results; that is, the Ss learning the related pairs turn in a superior performance as compared to the Ss

---

[4]In the binary notation the first seven digits convert as follows: 1 into 001, 2 into 010, 3 into 011, 4 into 100, 5 into 101, 6 into 110, and 7 into 111. Thus a string such as 101001111110011010 could be broken into groups of three, thus 101 001 111 110 011 010 and recoded as 5 1 7 6 3 2. The latter is in base eight and thus is called 'octal'.

learning the unrelated lists (Umemoto, 1962; Baddeley, 1964; Homzie and Weimer, 1967). A theoretical explanation of this might be that the related words possess a greater potential for the formation of effective mediational links. General research on mediation has shown that the number of trials, the number of errors, or the time to criterion are reduced and retention is improved when mediational links are employed (Wallace, Turner, and Perkins, 1957; Spiker, 1960; Jensen and Rohwer, 1963a & b). Furthermore, the degree of facilitation varies directly with the sophistication of the mediational strategy employed (Martin, Boersma, and Cox, 1965). Thus, although mediation is not in itself a form of organization it is, nevertheless, a very important determiner of the degree of facilitation that occurs as a result of organization. But, the degree of facilitation due to mediation or due to organization in general is, in turn, governed or determined by at least four factors. The first three are as follows: (1) the nature of the material, that is, its potential for organization or, as Bruner would say, its 'codability'; (2) the Ss' ability to perceive any potential inherent in the material or built into it by the researcher for a specific experimental reason; and (3) the Ss' ability to use the perceived potential for mediation, coding or organization.

That the nature of the material affects the degree of facilitation in P-A learning was evident from the studies by Umemoto (1962), Baddeley (1964), and Homzie and Weimer (1967). The same is true for other learning tasks. For example, Frincke (1968) found that in free recall tasks, words high in imagery-concreteness were more readily organized or clustered into groups than abstract, low imagery words;

and Lieberman and Culpepper (1965) found that Ss learning a list of objects did better than those learning a list of words for those objects.

Nevertheless, any amount of potential is useless if the Ss do not perceive it or if they do, fail to use it. At times there are alternative ways of organizing the material not all of which are equally facilitative, and sometimes the Ss do not choose the most effective form of organization. In one experiment by Tulving (1962b), half the Ss were instructed after the third trial to associate the words with the alphabet and to recall them that way on the test trial; the other half of the Ss were told that they were in the standard group and to do their best. The results showed that the alphabetical group did better with respect to mean number of words correctly recalled and by the eleventh trial many Ss recalled all words alphabetically without error. Tulving concluded:

> The experimentally produced difference between the two groups clearly was not that between alphabetical organization and no organization, but rather between alphabetical organization and other, apparently less effective form of organization (p. 189).

It is interesting to note that two out of about forty Ss in the standard group used alphabetical order in their recall while the rest did not. Thus it is rather obvious that given an equal organizational potential in the material the degree of facilitation will depend on the Ss' ability to perceive this potential, select the most effective organizational technique, and then use it successfully. That something like an 'organizational ability' may actually exist was indicated in a study by Earhard (1967) who was able to separate the Ss into 'high organizers' and 'poor organizers' and to show that these

two groups performed differently on a subjective organization task. Furthermore, a study by Jenkins and Russell (1952) showed that female Ss organized the material into clusters to a greater extent than male Ss and that they also recalled more words on the free recall test. Some of these individual differences might even be magnified further by learning or training. For example, Senter and Hauser (1968) showed that Ss trained in the commercial 'hook' or 'peg' system did significantly better than untrained Ss. The technique involves learning a set of verbal pegs and then the formation of 'bizarre images' so as to associate the test list with the previously overlearned 'pegs'. All this suggests that different degrees of facilitation may be obtained depending on the group of Ss that the experimenter happens to select.

The fourth determiner of the degree of facilitation is the opportunity that Ss are given to organize the material. If, for example, the Ss are given only a limited opportunity to organize the material then the degree of facilitation cannot be expected to be the same as when they are given extensive freedom to structure or arrange the material to their own pattern. Of the three most common learning tasks: P-A, serial, and free recall, the first gives the Ss the least amount of freedom. In a P-A task E determines which stimuli and responses are to be grouped into pairs; and any overall patterns or regularities are broken up by the common practice of presenting the pairs in a different random order on every trial. Thus any attempts at organization are frustrated and the Ss are faced only with the task of inventing mediational links or of using the ones built into the material by the experimenter. Furthermore, the perception of any

regularities or patterns of similarity are made difficult by the standard technique of presenting the pairs at a very rapid rate and one-at-a-time. Thus, the Ss are not only prevented from scanning the material as a whole for any inherent similarities or patterns, but they are also given a very limited time in which to discover and use any organizational strategies. The serial learning task offers just as little in terms of freedom to organize; however, it does retain some stability in that the material is presented in a constant order on all trials and must be learned in that order. Nevertheless, it is E who determines that order and not the Ss. The free recall task, in comparison, gives the Ss the most freedom, but even here the opportunity is limited by the same practice of using very short time intervals, randomizing the StOs after every trial, and presenting the material one StO at a time so that any existing patterns of regularity are not immediately apparent. However, the Ss do have the freedom to choose the order in which they recall the list on the trials and on the tests.

Some variations on these methods of presentation have given the Ss more freedom to organize while others have expanded the range of cues that are made available to the Ss. The latter have shown that the Ss are quite capable of using a wide range of attributes and cues to facilitate learning. One such set of experiments presented the StOs in a spatial array rather than in one common location. Results indicated that the presentation of nonsense syllables in a spatial pattern in front of the Ss, or that of real objects in specific locations on a table, facilitates learning (Asch, Hay, and Diamond, 1960; Berlyne, 1966). However, it was discovered that different configurations tended to have different degrees of facilitation (Mellgren, 1967).

In summary, then, it seems to be reasonably well established that (1) organization tends to facilitate learning; (2) different strategies or techniques result in different degrees of facilitation; (3) there are at least four interacting factors that govern this degree of facilitation; and (4) the standard methods of presenting the material to the Ss such as the P-A, serial, and free recall techniques do not give the Ss a sufficient opportunity to examine the material for possible organizational potential and then to use their ability so as to implement an organizational strategy. The lack of this opportunity and its effect on learning efficiency will be discussed in detail in the next chapter where similarity will be examined as one of the bases of organization.

Chapter 5

A POSSIBLE RESOLUTION OF THE SIMILARITY PARADOX

On the basis of the evidence presented under Propositions I, II, and III, and assuming that similarity is one phenomenon, it seems that similarity should facilitate learning.

In Chapter 4 the first section, Proposition I, showed that subjects (Ss) do not primarily learn by rote but, given a learning task, attempt to classify, code, or in general to organize the material into classes, concepts, patterns, or structures so as to facilitate learning. The second section, Proposition II, showed that similarity has great potential for organization, and that Ss form the habit of using similarity as a basis of classification and organization from childhood when they learn concepts or the words of a given language, and further, that this habit is carried over to learning tasks in general. The third section, Proposition III, indicated that any type of structuring or organization generally facilitates learning. Consequently it is reasonable to contend that similarity as a major source of organizational potential should facilitate all learning. However, as the section on the effects of similarity on learning showed, this is not always the case.

There are many possible reasons for this, but the most important one is that, if Ss are to use similarity as an organizational tool, then they must have the opportunity to do so. This means that

the Ss must have the time to examine all the material as a whole and not one stimulus object (StO) at a time. Furthermore, they must have the freedom to rearrange the material into a pattern or organizational scheme which is facilitative to their memory.

Nevertheless, the most common experimental techniques do not provide sufficient time for organization or even for the formation of effective mediational links. For example:

> In paired-associate learning the 'standard' rate of presentation is 2:2 sec., that is 2 sec. for the presentation of the stimulus and the anticipation of the response and 2 sec. for the presentation of the stimulus and response term together. If mediation of any degree of complexity is considered and is asserted to have a reality in the association processes of the S, it is quite clear that a 2-sec. anticipation interval is simply insufficient time for a chain of associations to be 'run off'. Yet, how many studies have increased the anticipation interval to allow for this? Very few, and yet...[some]...studies...conclude that if mediation does occur it has little or no effect on the rate of learning. How can mediation be expected to influence learning when the situation is such as not to allow time for mediation to occur (Underwood, 1964, p. 76).

Miller, Galanter, and Pribram (1960) describe a project carried out at the University of Pennsylvania by Wallace, Turner, and Perkins (1957) which showed that a person's capacity for the formation of associations is practically unlimited. The Ss were presented with pairs of English words and proceeded to form, at their own pace, visual images connecting the two words. Although the list of paired-associates was given only once, fantastic results were obtained. Up to 500 pairs, the Ss were remembering about ninety-nine percent; at 700 pairs this dropped to about finety-five percent. Initially the time that Ss took to form the associations was about twenty-five seconds, but as they became more experienced they could work with less than five seconds per pair. This, however, is more than twice as long as the 'standard' two second

interval during which a stimulus and response is available to the Ss in a paired-associate (P-A) task. Furthermore, in the project by Wallace et al. the Ss recorded their responses in writing which in many cases must have taken more than two seconds per item. However, extending the time interval in a P-A task may not be enough because the Ss would still have to work within a regular or equal time interval per item or pair. Thus they would still be under pressure to do and learn at a paced rate. This is very significant in view of the findings by Levitt and Goss (1961) which indicated that stress was facilitative with the low stimulus similarity lists and inhibitory with the high stimulus similarity lists (p. 248). If the latter effect is taken in conjunction with the relatively rapid pace of P-A learning, then it seems that the P-A technique unduly favours the learning of low stimulus similarity lists and places the Ss learning a high stimulus similarity list at a distinct disadvantage. As a result the negative effects attributed to similarity could be magnified out of proportion.

Although the above remarks are directed primarily at the P-A learning experiments the same conditions exist in serial and free recall techniques, since their presentation is also paced and uses equally short time intervals per item, thus making it very difficult for Ss to formulate and use any relatively effective organizational strategies. In addition to the inappropriate time factor all three techniques usually present the StOs to the subjects one StO at a time. Thus the Ss do not have the opportunity to examine all the material simultaneously for similarities, regularities, configurations, or patterns, etc. Consequently, even if these exist, the Ss may not be able to utilize them effectively assuming, of course, that they had

been perceived in the first place. Furthermore, any attempts at organization are made even more difficult by the common practice of randomization, that is, randomly rearranging the StOs after every presentation or trial. However, when the Ss are given the opportunity to examine all the StOs or, in other words, when the presentation is simultaneous and not one-at-a-time, then facilitation should result. Moursund and Chape (1966) found this to be indeed the case. Even under conditions or randomization they found that the Ss presented with all the pairs in the P-A task did significantly better than those who saw the pairs only one-at-a-time. Nevertheless, one cannot make a sweeping generalization that the whole (W) method of presentation should be superior to a one-at-a-time method of presentation, since the degree of facilitation in the former depends on the degree of organizational potential in the material and/or the Ss' ability to use that potential. In other words, if the organizational potential is low or if the Ss are lacking in organizational ability, then there may, in fact, be very little advantage to a W-presentation. For example, Brown and Read (1966), using 12-pair lists of CVC's low in formal similarity and of 75 to 85% Archer-association value, found that there were no significant differences between the two methods of presentation. This is not at all surprising since a close examination of the P-A list used failed to produce even a hint of any effective organizational scheme that the Ss might have used under conditions of W-presentation. Thus, one could not expect to find any facilitation due to organization, and therefore due to W-presentation. In fact, at times the W-method of presentation may even retard learning. If the Ss, for example, looking for means to organize the material, do not find any very effective

strategies either because there are none to be found or because they lack in ability, then they have, in fact, 'wasted' some of their time to no purpose. In the meantime, the Ss in the one-at-a-time group, because of the method, have not been searching as actively and in this case uselessly for strategies. In other words, the one-at-a-time method forces the Ss to learn the StOs as they come up, and any search for strategies has to be relegated to odd 'if-and-when' moments. Of course some incidental learning will occur in the first group but possibly not enough. Thus, the effectiveness of the W-presentation method depends on the organizational potential of the material and/or the Ss ability to use it.

An examination of the P-A list used by Moursund and Chape (1966) showed distinct organizational potential. For example, two out of ten P-A's ended and started with the same letter (DAX-XAP); six out of ten pairs had an initial letter in the stimulus and in the response member that were not more than four letters apart in the alphabet; one pair was in direct alphabetical sequence (YAV-ZAJ) while another was in reverse order (WEQ-VUD); and two pairs were very similar in sound (XIF-CEF) and had the same last letter. On the other hand, the list used by Brown and Read (1966) had only two pairs out of twelve in alphabetical order (JIB-KOY). The other pairs had initial stimulus and response letters at least three letters apart alphabetically and some were as far as sixteen letters apart (FET-WUN). Furthermore, half the pairs were in reverse alphabetical order (REM-BIF). Consequently, there would be little advantage to the W-method. However, in the former case a method which permits scanning, detailed examination, and quick double-checking without having to wait for the next trial does have distinct advantages.

In the Moursund and Chape list the Ss could have used a number of organizational, or in this case mediational, rules (e.g. 'stimulus and response sound similar' or 'last letter of stimulus and first letter of response is the same'). Furthermore, each rule had in most cases two or more exemplars, thus spotting and clearly differentiating the two exemplars and then grouping them under the same rule would clearly be easier under a W-presentation than under a one-at-a-time presentation. In the latter type of presentation the Ss might not be able to spot all possible grouping rules at once; but even if they did, they would still have to wait for the next trial before they could confirm their observations. It is not likely that any S after one presentation by the one-at-a-time method could have the four rules and their exemplars clearly in mind. On the other hand, a S in a W-presentation group could easily check and double-check the groupings and exemplars just by flicking his eyes back and forth over the list; that is, the S does not have to wait for the next trial to confirm any tentative or vague grouping ideas.

In free recall learning the material is usually of a meaningful nature and thus should have a relatively high organizational potential. Puff and Bousfield (1967), comparing the degree of organization, found that simultaneous or "...array presentation resulted in appreciably more (an average of three times as much) intertrial organization..." (p. 215). In view of the fact that "...there is a positive correlation between organization and performance" (Tulving, 1962a, p. 354), one can infer that W-presentation should also facilitate free recall learning. However, at present, the author is not aware of any studies that have examined this particular aspect of free recall learning.

As far as serial learning is concerned, a thesis by Keenan (1967),

comparing the effectiveness of six methods of presentation, found that the W-method resulted in significantly superior learning. The material used was a set of 'common three-letter words' and so should have a reasonable degree of organizational potential. Consequently the results are not at all unexpected. However, what is surprising is that so little has been done with the W-method of presentation particularly since it

> ...most closely approximates types of learning situations met by the student in actual school experience, and thus the explication of any differences in learning emerging as a result of this presentation method might be useful in attempting to generalize present ... findings to the school setting (Moursund and Chape, 1966, p. 167).

A similar sentiment had been voiced by Postman and Egan as early as 1949, but it did not seem to have had much impact on the type of research that was subsequently conducted.

Continuing with the original line of reasoning that organizational potential plus opprotunity to organize leads to a degree of facilitation, it would seem reasonable to conclude that a constant order of presentation should lead to better results than one in which the material is randomized after every trial. Again, however, one cannot make such a sweeping generalization because the effectiveness of a constant order of presentation depends not only on a reasonably high degree of similarity or organizational potential in general, plus the _Ss_' ability, but also on the specific order chosen for a specific list. That is, assuming both a high degree of organizational potential and able _Ss_, a given constant order may not lead to any better results than regular randomization after every trial _if_ the chosen order happens to be inappropriate to the potential inherent in the material. Earhard (1967) found that not only

could Ss be grouped into 'high organizers' and 'poor organizers' but that list sequences could be made in a 'preferred order' and in a 'non-preferred order', and that "...the preferred order of presentation resulted in superior performance for both types or organizers although HSO [high subjective organization] learners achieved superior performance to LSO [low subjective organization] learners within each order-of-presentation category" (p. 504). Earhard (1967) based her experiment on the indications found by Tulving (1962a) that under conditions of free recall the order in which Ss choose to recall lists of unrelated words do not seem to be entirely idiosyncratic and that there seems to be some commonality among Ss as to the order in which they recall the seemingly unrelated words. The preferred order list was, therefore, constructed from sequences of pairs of words most frequently found in the recall records of the sixteenth trial of Tulving's 1962 experiment, and the non-preferred order list was constructed from a sequence of pairs of words never found in succession in these recall records. Thus,

> ...whatever the basis for the final structure imposed by the previous learners - whether they had clustered on the basis of mediating responses, interitem associations, similarity in sound, familiarity, grammatical linkage, imagery - the structure or sequence of recall those learners had developed facilitated memorization, not only when explicit instructions to adopt the order of presentation were given, but also when Ss were free to adopt any sequential organization they chose ....
>
> The results indicated fairly consistently that it was easier to remember words which were presented in an order of recall preferred by previous learners (Earhard, 1962, p. 505).

The need to consider 'preferred order' and potential for organization as variables becomes even more apparent when the results of P-A experiments comparing a constant order and a changing order of presentation as well as poorotunity to organize (time variable) are considered as controlling factors then the results become less inconsist-

ent. At times, however, it is difficult to pinpoint the exact cause of some results since the authors very often give only a description of the StOs used to make up the list and, at other times, although the lists are quoted in full, the exact constant order used is not specified.

All six experiments to be desdribed used P-A pairs consisting of a nonsense syllable (trigram) stimulus and a meaningful word response. The method of presentation was either the 'recall-5' (r-5) or the 'recall-1-2-...-N' (r-1-...N). The first entailed five presentations followed by a test trial, the second, as indicated by the numbers had a test trial after every presentation for N trials. Some of the experiments also used a variation of the recall-5 method by adding on more presentations and tests; thus there are some experiments that presented the material to the Ss eight times before having a test trial (r-8), while others followed the test trial with further presentations and tests (e.g. r-5-10, or r-5-10-15). The general findings were that the constant order was either superior to the random order (Samuels and Jeffrey, 1968; Newman, 1966; Newman and Saltz, 1962; McGeoch and McKinney, 1937), or that there were no significant differences between the two methods of presentation as far as the mean number of correct responses was concerned (Martin and Saltz, 1963; Carluccio and Crowder, 1966). However, in comparing the list used by Samuels and Jeffrey (1968) with those used by Martin and Saltz (1963) it becomes evident that the former had stimulus terms high in formal similarity while the latter had stimulus terms low in formal similarity. The response terms in both cases were the same.[1] Therefore it is evident that under conditions of

---

[1]It is certain that identical response terms were used by Samuels and Jeffrey (1968), Newman (1966), Newman and Saltz (1962) and Martin and

one-at-a-time presentation, maintaining a constant order of items in the list from trial to trial seems to provide at least some opportunity for organization that is not provided when the items are randomized after every trial, and that this opportunity can only lead to facilitation if there is some similarity or organizational potential built into the list. Newman and Saltz (1962) did not quote the list that was used but from the description it is obvious that it had stimulus terms high in formal similarity. The results of an r-5 presentation showed facilitation under constant order. A replication of this study by Newman (1966) using the r-5-10-15 sequence did not confirm significant effects on the first test trial following the fifth presentation, and it was only on r-10 and r-15 that positive effects due to a constant order became large enough for significance. Since Newman did not quote the list one can only speculate that either the Ss were less able organizers or if a different constant order was used that it was not as 'preferred' as the previous one.

However, the general implication seems to be that the constant order is superior to the changing order of presentation only if there is some organizational potential built into the material. Nevertheless, a study by Carluccio and Crowder (1966) does not seem to support this conclusion. They examined the order variable under two conditions of formal stimulus similarity and found that although low formal stimulus similarity resulted in superior performance, the effects due to order

---

Saltz (1963). Furthermore from the description given by Carluccio and Crowder (1966) it is very likely that they also used the same set of response terms. Consequently, any variations in the results can be mostly attributed to variations in stimulus terms, and to variations in technique if any.

were not statistically significant. In other words, a constant order did not yield better results under conditions of seemingly high organizational potential (high formal stimulus similarity). One reason, of course, might be that the constant order selected did not happen to be the 'preferred' order. A second reason, and perhaps the obvious one might be the relatively fast test rate of two seconds per item used by the experimenters as compared to a rate of ten seconds per item used by all the researchers who obtained significant effects, particularly since test rate had been shown to be a significant variable (Newman, 1966). In view of the previously cited findings by Wallace, Turner, and Perkins (1957) that an appreciable time interval is required for the formation of effective mediational links, it seems quite probable that the lack of a significant order effect in the Carluccio and Crowder (1966) study may be attributed to the very short time interval used by the researchers. Thus the evidence seems to indicate that order is not a significant variable with lists <u>low</u> in organizational potential (Martin and Saltz, 1963; Carluccio and Crowder, 1966), but is significant with lists <u>high</u> in organizational potential (Samuels and Jeffrey, 1968; Newman, 1966; Newman and Saltz, 1962; McGeoch and McKinney, 1937) unless the rapid rate of presentation makes it impossible for the <u>Ss</u> to use any effective organizational strategies, thereby turning the organizational potential of the material into an <u>irrelevant</u> factor.

All the P-A studies cited above used a nonsense stimulus term and a meaningful response term. When both members of the pair are meaningful words one can expect that the order effect would be rather large. This is indeed the case. For example, a study by McGeoch and

and Underwood (1943) using a four seconds per item presentation rate found that the Ss in the random condition required one and two thirds times as many trials to reach the criterion of two perfect recalls as the Ss in the constant order condition. Furthermore, the superiority of the latter was apparent after the fourth trial. The four second interval may not seem to be very long; nevertheless, it is twice that used by Carluccio and Crowder (1966) and it was used with meaningful material which is perceived and apprehended much more readily than nonsense material (Miller, Bruner, and Postman, 1954). Consequently the results are not surprising. If nonsense syllables are used for both members of the pair then one may expect that a longer time interval would be needed in order to obtain significant order effects. Rubin and Brown (1967), using the mean total errors as a measure of learning and a five second time interval in the r-1-...-N presentation pattern, found that order was indeed a very significant variable.[2] Furthermore, the study showed that random presentation of pairs as compared to constant presentation unduly favoured the low formal similarity lists (low stimulus similarity-low response similarity pairs). The high stimulus similarity-high response similarity list had an increase of 113.9% in the number of errors when the order was random rather than constant. On the other hand, the low-low lists had 5.6% FEWER errors when the order was random instead of constant. Furthermore, the spread in errors increased anywhere from 5.5 to 8 fold for

---

[2] It is quite possible that different measures of learning as well as different presentation techniques (e.g. r-1-...-N versus r-8) are not equally sensitive to the order variable. Consequently, the exact relationship between time interval and significant effect may depend on more than the nature of the material.

variations in stimulus similarity and only twofold for variations in response similarity when the order of presentation was random rather than constant. Thus, from this study it is evident that stimulus similarity lists seem to be affected much more by the order variable than response similarity lists, and that any high formal similarity list is affected much more by randomization than any low formal similarity list.

Summarizing all the above findings it seems that very often the Ss do not have an adequate opportunity to use similarity as a basis of organization because of the one-at-a-time method of presentation, the relatively short time intervals used, and the very common practice of randomizing the StOs after every trial. Also as a result of these techniques the Ss are placed under stress which seems to facilitate the learning of low formal similarity lists and to retard the learning of high formal similarity lists. Therefore, the effect that similarity has on learning depends on three factors: the extent to which Ss are given the opportunity to use similarity for organization, their ability to perceive this potential, and their facility in using it once it is perceived. However, since the opportunity to use similarity depends on the learning conditions and since these vary from experiment to experiment within such broad techniques as P-A, serial, and free recall, conflicting findings are to be expected within each broad class of learning techniques. In other words, some of the contradictory results described in the section on the effects of similarity on learning may be explained in terms of an interaction between two variables; namely, similarity and the learning conditions. However, this will not by any means explain all the contradictory

findings. There are a number of other very important considerations that have to be clarified.

A quick examination of the previous section on the effects of similarity on learning will show that whereas meaningful similarity generally does not hinder learning, and in many cases seems to facilitate it, formal similarity generally hinders learning. The different effects of similarity on learning might, of course, be attributed to the factors summarized above; however, even under the most favourable conditions it is quite probable that high formal similarity would not result in a better performance. The reason for this is that high formal similarity does not guarantee high organizational potential, and it is the latter which leads to facilitation under conditions of opportunity.

It is quite possible to construct two lists equal in formal similarity but having different degrees of organizational potential. For example, the two lists in Figure 1 are constructed in accordance with the description given by Carluccio and Crowder (1966); however, as a close examination will show, List 1 has a greater organizational potential than List 2. Both lists are composed of "... CVC nonsense syllables containing only the letters X, V, F, Z, U, and Y ..." and ranging in " ... Glaze association values ... from 0% to 20% (Underwood and Schulz, 1960)..." (Carluccio and Crowder, 1966, p. 615). List 1 contains six CVC's which are all the possible combinations of V with X, Z, and F using the two letters U and Y as vowels; and four inverses to make a total of ten syllables. List 2, on the other hand, is seen to use only two of the possible three combinations in each case (VYX, VYZ, but not VYF), one inverse of a CVC that is in the

List 1

VYX	(and its reverse)	XYV
VYZ	"	ZYV
VYF	"	FYV
VUX	"	XUV
VUZ		
VUF		

List 2

VYX	(and its reverse)	XYV
VYZ		-
-		FYV
-		XUV
VUZ		-
VUF		-
XYF		
ZYF		
ZUX		

Figure 1. Lists equal in formal similarity but differing in organizational potential.

list, two inverses of CVC's that are not in the list, and three combinations that do not fit the two 'generating rules'. The six 'V' trigrams of List 1 make up what is called an exhaustive category. Cohen (1963) showed that recall is significantly better for exhaustive category StOs than for non-exhaustive category items. Although Cohen used meaningful words and exhaustive categories such as north, south, east, and west, it is quite possible that a similar effect may manifest itself with CVC's, particularly since a study by Segal and Halwes (1965) showed that Ss are quite capable of discovering and using the generating rules by means of which the experimenters constructed the nonsense syllables in the first place.

Another way of describing the difference between the two lists would be to say that List 2 has a greater degree of chaos or randomness, while List 1 is much more orderly because it is constructed according to certain rules which create sequential constraints among the letters of the trigrams. A study by Miller (1958) examined the free recall learning of letter strings which were generated either randomly or by rules producing sequential constraints and found that the latter resulted in significantly superior performance. Garner (1962) cites many other studies which support Miller's findings.

Another difference between the lists is that List 1 requires less learning than List 2. For example, the Ss learning the first list can recode the material into a number of generating rules such as follows:

(1) Form all combinations of VY with X, Z, and F
(2) Take all the mirror images of the above.
(3) Form all combinations of VU with X, Z, and F.
(4) Take only the X inverse from the last set.

Alternately, the Ss may use a more mathematical approach:

(1) Form all combinations such that:
  (a) the first letter is V,
  (b) the second letter is Y or U,
  (c) the third letter is X, Z or F.
(2) Form all the inverses of the VY combination.
(3) Form only the X inverse in the VU combination.

However, when it comes to List 2 there is very little recoding possible in terms of generating rules. Thus, while List 1 can be mastered by learning three or four general rules, List 2 can only be mastered by learning approximately ten separate CVC's which are composed of common elements that are combined in what is effectively a random way.

Another way of characterizing the difference between the lists is to consider organization in terms of similarity based grouping. Thus the items in List 1 may be grouped into three sets: the VY items and their mirror-images, the VU items, and the single XUV trigram. The items in List 2, however, cannot be grouped into a 'few neat groups' or 'chunks' but require much more complex attempts at organization. Furthermore, whereas the groups in List 1 were mutually exclusive, the groups that are possible for List 2 would be overlapping, that is, an item could belong to more than one group.

Taking all these factors into account it is not likely that anyone would disagree with the statement that List 1 would be the easiest one to learn. However, the exact degree of superiority in performance would depend to a large extent on the sequential ordering of the trigrams in the list. If, for example, the items of List 1 were presented in an order which masked the generating rules or the possible groupings, then the learning of the list would be slowed down Also, it now becomes rather obvious why one-at-a-time presentation,

short time intervals per item, and constant randomization have such retarding effects on the learning of lists that are high in organizational potential (and high in formal similarity). Under such unfavourable conditions the perception of the potential is slowed down. In time, of course, some of the Ss will perceive the generating rules or the possible groupings in spite of these methodological obstructions; however, other Ss will never perceive them. Thus, facilitation due to organizational potential may occur, may be delayed, or at times may be completely eliminated. The learning of type 2 lists is also retarded by the same presentation techniques since the Ss do not have adequate time for proper discrimination; however, the variability is much less with lists of this type. Consequently, lists differing in organizational potential under the same learning conditions can lead to different results even if they are equated on formal similarity.

This difference in performance has been shown to exist in free recall studies where 'form of structure' was varied while formal similarity was held constant. It was found that formal similarity hindered free recall only with items that had poor form of structure (Garner and Whitman, 1965; Whitman, 1966; Koeppel, 1967 and 1968). Thus two lists of equally high formal similarity were learned at different rates depending on the goodness of form. This is something which had never before been taken into account.

The idea of form of structure originated with Garner (1962) and was later developed by his associates. Essentially the form of structure is considered to be good if there is a definite relationship between the elements in different positions; or in other words, if

"... a high correlation exists between specified letter positions" (Koeppel, 1968). A thesis by Koeppel (1967) illustrates the difference between good and poor form. If the numbers 1, 2, and 3 are used to make up trigrams, then Set 1 in Figure 2 has good form of structure and Set 2 has poor form. Good form in this case is obtained by having a 1 in the first position always followed by a 1 in the second position. Poor form is achieved by following the 1 in the first position either by a 1, 2, or 3. In other words the correlation between letter positions in the case of poor form is low. In trigrams, of course, letters are not duplicated within a given nonsense syllable, but the same sort of sequential constraints are possible.

Comparable results were obtained in P-A studies where relatively restrictive learning conditions operate. In one study Runquist (1968a) found that by varying the 'heterogeneity' and 'variability' properties of lists which were equal in formal stimulus similarity, different performances could be obtained. 'Variability' refers to the different degrees of pair-wise similarity that are possible when every stimulus item in the list is compared to every other stimulus item. Thus, in some lists all the items may virtually be equally similar to each other (low variability); in other lists, some of the items may be very similar while others are dissimilar (high variability). 'Heterogeneity', on the other hand, refers to the number of different rules describing the locus of similarity. Runquist (1968b) identified 23 such rules for CCC's and 11 for CVC's. Each rule describes the number of common letters in a pair and its position. Table 1 lists these rules for CVC's. Although the rules differentiate between the number of common elements in a pair, the property of 'heterogeneity' refers

<u>Set 1</u>

111	221	331
112	222	332
113	223	333

<u>Set 2</u>

111	213	312
123	222	321
132	231	333

Figure 2. Lists Illustrating Good (1) and Poor (2) Form.

TABLE 1

RULES FOR LETTER-POSITION IDENTITY IN CVC PAIRS[a]

Group	Rule	Example
0	No repeated letters	XUQ-PIJ
1A	First letter repeated	XUQ-XAH
1B	Vowel repeated	XUQ-TUV
1C	Third letter repeated	XUQ-YEQ
1D	First letter in third position	XUQ-YOX
2A	First two letters repeated	XUQ-XUL
2B	Second two letters repeated	XUQ-WUQ
2C	First and third letters repeated	XUQ-XOQ
2D	First letter in third position. Vowel identical	XUQ-VUX
2E	First and third letters reversed. Vowel different	XUQ-QOX
3	Spelled backwards	XUQ-QUX

[a]Taken from Runquist and Joinson, 1968, p.317.

only to the number of different rules occuring in a list.

The study indicated that a list high in variability and low in heterogeneity (that is, homogeneous) produced performance which was superior to that on the low similarity list; however, the difference was not statistically significant. Consequently, even under the restrictive P-A learning conditions formal similarity lists may be so constructed as to yield performance which, although not better than that of the low similarity lists, is not significantly different from it. Runquist (1968a) concluded that

> ...the relationship between formal intralist stimulus similarity and performance cannot be characterized in terms of a simple function. Considerable modulation of performance has been obtained among sets of stimuli in which overall [formal] similarity was virtually the same (p. 640).

Similar results were obtained with CCC's where lists with 18, 13, and 8 different letters did not yield significant differences in performance while manipulation of the number of rules (heterogeneity) produced significantly different performances on lists with an equal number of common letters. In general, therefore, in P-A tasks the indication is that "...letter identity per se is not a very important variable with CCC stimuli" (Runquist, 1970a, p. 158).

In all P-A tasks there is one additional factor peculiar to that learning task which may affect performance differentially. The Ss learning the low stimulus similarity lists have one very important advantage: all that they have to do is to associate the first letter of the stimulus with the response. They do not have to learn the other two letters because in a low formal stimulus similarity list all the first letters of the nonsense syllables are usually different. However, the Ss learning lists with stimulus terms high in formal

similarity cannot do so since many letters are repeated. Consequently, they have to learn either two letters per stimulus or at times all three letters. This added amount of learning, although minimal, will nevertheless slow them down. That stimulus selection does occur has been documented under Proposition I (p. 57). Nevertheless, one qualification is necessary. Runquist (1968a) found that stimulus terms which had identical vowels and last letters did not lead to a superior or even equal performance on a P-A task as compared to a low stimulus similarity list. This seems to suggest that Ss pay less attention to the actual letter structure than may be expected and that the stimuli are coded as pronounceable units rather than as a collection of letters (Runquist, 1968a, 1970b). Nevertheless, it cannot be assumed that all the Ss pronounce all or even a large percentage of the StOs. however, the pronounceability of the items in the list is another factor which may lead to different results with seemingly comparable lists or stimuli.

In brief summary, it is possible, therefore to construct lists equal in formal similarity but differing with respect to such variables as form, heterogeneity, and variability, and thereby to obtain different performances. As the illustration in Figure 1 (p. 94) showed, 'one' list description can result in two very different lists which would lead to different results under the same learning conditions.

Thus, it is not at all surprising that studies using formal similarity as the only variable produce contradictory results within each of the three common learning techniques. An examination of the lists used in verbal learning studies showed that many of them were

like List 2, Figure 1, p. 94, some were like List 1, while others were somewhere in between these two examples or beyond them at one of the extremes. For example, a list such as ZGJ, GQJ, JFQ, QFZ, FQJ (Underwood, 1966, p. 476) is very high in formal similarity, much more so than List 2. However, since the specific relationships between the letters and among the StOs of a list have not been regarded as relevant until very recently, many of the researchers did not quote the exact lists but rather gave short enumerations of the letters used. Consequently, in such cases, it is impossible to determine if the negative effects on learning occur because of a lack of opportunity to use an existing organizational potential, that is, the confounding of methodological factors with organizational potential, or because the lists were randomly generated so that there was no organizational potential available in the first place.

The joint effect of the three variables of formal similarity, organizational potential, and learning conditions is summarized in Table 2. The entries in the six cells indicate the possible performance under conditions of interaction between two levels of formal similarity (fS), two levels or organizational potential (OP), and two levels of learning conditions: 'favourable' and 'unfavourable'. By 'favourable conditions' it is meant, presentation of all StOs at once, adequate time for examination of the list, and the maintenance of a constant ordering of the items from trial to trial. On the other hand, by 'unfavourable conditions' it is meant, one-at-a-time presentation of StOs, very short time intervals per item, and randomization of the sequence of StOs after every trial. Between these two extremes there are, of course, a number of variations, particularly with respect to

TABLE 2

THE JOINT EFFECT OF FORMAL SIMILARITY (fS), ORGANIZATIONAL POTENTIAL (OP), AND LEARNING CONDITIONS ON PERFORMANCE

	High Formal Similarity		Low fS
	High OP	Low OP	
Favourable Presentation Conditions	Very Good	Poor	Fair - Good
Unfavourable Presentation Conditions	Fair	Very Poor	Fair -- Good

the length of the time interval either per item or per list. Also there are different degrees of OP as well as different degrees of formal similarity that can be built into the lists. Thus the table shows only the possible performance for six possible combinations out of an almost unlimited number of interactions. Is there little wonder, therefore, that comparisons between performance of high formal similarity (HfS) lists and low formal similarity (LfS) lists did not yield consistent results? For example, if OP happened to be high (HOP) and the presentation conditions were very favourable (perhaps as a result of a longer time interval and/or a constant order of presentation) then it would be possible to get results which indicated that HfS facilitated learning. However, if the list was low in OP (LOP) and the conditions were extremely unfavourable (which is true in most of the studies) then it would be possible to get results which showed that HfS markedly retarded learning. And, under conditions of HOP and somewhat unfavourable presentation conditions, it would be quite possible to get results which showed that HfS was neither facilitative nor retarding in its effect on learning. All of these results have been reported in the literature at one time or another (see pp. 24-32).

The only solution seems to be the incorporation of OP as <u>another variable which has to be considered whenever formal similarity is studied</u>. However, this is <u>not</u> an adequate solution because it would still leave unsolved the 'meaningful-formal' similarity contradiction. Similarity is <u>one</u> phenomenon whether formal or meaningful, and therefore it should affect learning in the same way under comparable conditions. Nevertheless, even under the most favourable presentation conditions it is highly unlikely that HfS would lead to

facilitation as long as OP was low. Consequently, this would still contradict the findings that high meaningful similarity generally facilitates learning.[3] Thus, taking OP into account still does not solve the problem of two different effects of similarity on learning; and using meaningful similarity, formal similarity and OP as three distinct variables is not scientifically parsimonious nor is it consistent with the idea of one similarity phenomenon.

Since the contradictory effects can be narrowed down to the different roles played by high formal-LOP-similarity and meaningful similarity, an examination of these two types of similarity is in order. One very basic difference becomes immediately apparent; whereas meaningful similarity can be used for grouping like StOs together, formal-LOP-similarity cannot be used for this purpose.[4] In fact, the only time that formal similarity can be used to group or classify like StOs is when it has a HIGH degree of OP. On the other hand, meaningful similarity does not have this problem since it always has OP, and thus can always be used to group StOs into classes or sub-classes. It becomes apparent, therefore, that OP is a property of meaningful similarity. Consequently, the only possible conclusion consistent with the one-phenomenon idea is that since high formal-LOP-similarity does not

---

[3] Any non-facilitative effects of meaningful similarity on learning may be, with a high degree of certainty, attributed to unfavourable presentation conditions.

[4] More precisely, the difference is that meaningful similarity can be used to form a few classes composed of many items; formal-IOP-similarity can only be used to form very many classes composed of very few items and very often there is only one item in the class, or at most two. Thus grouping does not provide much saving as far as learning is concerned.

possess this property it is *NOT similarity at all* (or more precisely, the similarity is so low as to be negligible). This somewhat startling conclusion becomes very obvious if StOs other than nonsense syllables are considered. For example, if the following letter matrices (two dimensional nonsense syllables) are examined, it will be seen that while matrix 1 and 2 are very similar, matrix 3 is very different from either matrix 1 or 2.

```
A B A B A      B A B A B      D C A C A
B C D C B      A D C D A      E D A A B
A D E D A      B C E C B      A D B B B
B C D C B      A D C D A      C D A A B
A B A B A      B A B A B      B B A B C
   (1)            (2)            (3)
```

Matrix 1 and matrix 2 both have a regular relationship among the letters; matrix 3, however, was constructed by using a table of random numbers and thus there is no meaningful relationship among the letters. Matrix 1 and 2 are similar because they stand in a very definite and regular relationship to each other. This relationship is missing between matrix 3 and the other two matrices. However, matrix 1 and 2 have more than just a sequential relationship among the letters. Matrix 1 and 2 also have a *specific* relationship between the letters and the matrix coordinates. Thus A's and B's occur only in the border positions, while E occurs only in the central position. The letters C and D form only the inner framework. Had these letters possessed only sequential relationships then the matrix might have looked as follows:

```
E D A B A
B A B C D
C B A B A
B C B C D
A D A B A
   (4)
```

Matrix 4 has exactly the same sequential constraints as matrix 1. That

is A is followed only by B or D; B is followed only by A or C; C is followed only by B or D; D is followed only by A, C, or E; and E is followed only by D. A comparison of the degree of similarity between matrix 1 and 2, and between matrix 1 and 4 will show that whereas the similarity between matrix 1 and 2 is very high, that between 1 and 4 is not high at all. The reason for this difference, of course, is that matrix 1 and 4 share only common sequential constraints; while matrix 1 and 2, on the other hand, share not only almost identical sequential constraints, but <u>also</u> the same relationship between the letters and the coordinates of the matrix. Thus for a <u>high</u> degree of similarity to exist between two matrices there must be a sequential relationship between one letter and the next (sequential constraint), AND a relationship between a given letter and its spatial coordinates or location in the existing frame of reference, as well as some common elements. Therefore, the degree of similarity that exists between matrix 1 and 2 is much greater than any similarity that migh exist between matrix 3 and another randomly constructed matrix, because all that the latter pair would have in common would be a number of identical elements (letters), while the former pair have the same number of identical elements as well as a set of common sequential and positional relationships. In comparison the degree of similarity between two random matrices is quite insignificant and thus cannot be termed 'high'. Nevertheless, according to the formal similarity definition, the degree of similarity between the two randomly constructed matrices would have to be considered as being high. Obviously this is not the case. Formal similarity, as it is defined, does not take into account <u>anything</u> except the <u>number</u> of common elements and thus rates two StOs as

being highly similar when in actual fact they are not very similar at all.

In summary, all four matrices were constructed from five different letters which were used equally often in the matrices, thus, the formal similarity among them was very high. Nevertheless, the similarity between matrix 1 and 2 was actually vastly higher than that between 1 and 3, 2 and 3, 3 and another random matrix, 1 and 4, 2 and 4, or 3 and 4. If two lists were constructed so that one contained matrix 1 and 2 as well as others possessing the same characteristics, then such a list would be higher in OP than one containing randomly constructed matrices such as 3 as well as some which possessed only sequential constraints such as matrix 4. While in the former list the matrices could be grouped according to the different types of spatial relationships and sequential constraints, the latter list could be said to consist of effectively different or, at best, low similarity matrices, since no two could be put into a category with another matrix even though they all used a common set of elements. Consequently, the latter list would have to be classified as a HIGH formal-low organizational potential-similarity list which was composed of essentially DIFFERENT StOs, or at best, StOs which were actually LOW in similarity. Thus, what seems to be needed is a new definition of similarity for nonsense StOs which would rate StOs as highly similar only if they were also high in OP. Such a definition would then be consistent with the relationship that exists between meaningful similarity and OP, and the idea of one similarity phenomenon. In fact, the need to eliminate the combination of high formal similarity and low OP is a direct consequence of the 'one phenomenon assumption' because if similarity is one

phenomenon, then not only must similarity affect learning in the same way regardless of what type of StOs are being used, but the StOs classified by a similarity definition into high, medium and low must exhibit parallel characteristics and posses parallel properties regardless of the nature of the material. Since similarity is highly correlated with OP and serves as a basis for organization when the StOs are meaningful items, then the same should hold when the StOs are nonsense items. In this respect the formal similarity definition is inadequate because it omits two whole dimensions: sequential constraints and spatial coordinates.

These two dimensions can be collectively referred to as the 'internal organization' of the StO. And, just as StOs may be composed of common elements so they may also have like internal organizations or different internal organizations, and the degree of similarity between two StOs would depend on the closeness between their individual internal organizational schemes as well as on the number of common elements and their role in the overall configuration.

Of the three factors entering into the new similarity definition, the relationship between the elements and the framework of the StO is the most important and influential factor in determining the degree of similarity. For example, two StOs may be similar (although in a low degree) and yet possess no common elements whatsoever. The following two matrices illustrate this possibility:

$$\begin{array}{ccccc} O & C & O & C & O \\ C & V & W & V & C \\ O & W & S & W & O \\ C & V & W & V & C \\ O & C & O & C & O \end{array} \qquad \begin{array}{ccccc} H & N & H & N & H \\ N & I & J & I & N \\ H & J & X & J & H \\ N & I & J & I & N \\ H & N & H & N & H \end{array}$$

(5) (6)

In spite of the fact that completely different elements are used to make up the configurations, the pattern stands out and makes the matrices similar to some extent. This similarity is much greater than that between matrix 3 and 4 which possess no common internal organization but only a number of common elements.

The inadequacy of the formal similarity definition, or more generally the common elements definition, is more evident with three dimensional StOs. Suppose two chairs, A and B, are taken as the StOs such that they are identical except for the arm rests. All would agree that they are highly similar. However, if one of the chairs is taken apart and then put together again but in a completely different way (different internal organization) then the degree of similarity between an untouched chair, B, and the reassembled chair, A', would be very low. Nevertheless, the formal similarity between B and A' would still have to be given as 'high'. It seems, therefore, that again <u>the formal similarity definition leads to ratings which are at variance with observation</u> or even painstaking analysis.

The same arguments, of course, apply to the similarity between nonsense syllables, but here the importance of the sequential constraints and particularly that of the spatial coordinates is less self-evident because the StOs are linear or one dimensional. Nevertheless, the internal organization of a StO is just as important. For example, the StOs in a given Set A, where A = [KXYS, KXYO], are highly similar, but the same cannot be said for the items in Set B, where B = [KXYS, YKOX]. In Set A the items have both identical elements in the same sequence and at the same spatial coordinates; in Set B the letters possess neither common coordinates nor the same sequence. A set such

as [KXYS, OKXY] illustrates two StOs which are similar because they have a common sequence of items; a set such as [KXYS, KSYX], on the other hand, shows two StOs which are similar because they have two common elements in the same locations. The highest degree of similarity exists between the StOs of Set A, the lowest between the StOs of Set B; and an intermediate degree of similarity exists in the other two sets.

If a definition of similarity is to differentiate among all these degrees of similarity, then it obviously must make some reference as to the internal organization of the StOs. That the spatial coordinates of the letters in trigrams are relevant has been implied by a limited number of studies; however, none of these attacked the question of the similarity definition directly (Martin, 1968; Nodine and Hardt, 1968; Monty, Fisher, and Karsh, 1967; Runquist and Joinson, 1968). The experiment by Martin (1968) indicated that Ss do seem to attach particular significance to the first letters of the trigrams, and that it was possible, by manipulating the letters in this position, to influence learning quite significantly. The study by Nodine and Hardt (1968) examined the 'Archer meaningfulness values' for trigrams and letters and found that the 'meaningfulness' of a letter changed with its position in a trigram. A third study, even more marginally related to the internal organization of StOs, was conducted by Monty, Fisher, and Karsh (1967). They found that in tallying letters Ss reported using spatial cues by visualizing a number of windows in which the tallies were kept, that is, one window per letter. Finally, there is the study by Runquist and Joinson (1968) in which Ss judged the similarity of pairs of trigrams. The CVC's differed in 11 ways depending on the

location of one or two common letters (see Table 1, p.100). The results indicated that the spatial coordinates of the common letters significantly affected the similarity judgements. Since these results were based on judgements and not an analysis the findings pertain more properly to psychological or subjective similarity than to physical or objective similarity. Nevertheless, the results may be taken to imply that spatial coordinates of letters should be examined as possibly valid <u>characteristics to be incorporated in a definition of objective similarity</u> for nonsense StOs.[5]

All the arguments presented in this Chapter have already illustrated this need, not only as far as spatial coordinates are concerned, but also for sequential constraints. Both aspects are needed for a <u>comprehensive</u> definition of objective similarity for nonsense materials.

The new definition which is being proposed would apply to all physical StOs with the exception of printed meaningful words; and might be stated as follows:

> The degree of objective similarity varies directly as the number of common elements, and/or the number of common sequential constraints, and/or the number of common spatial relationships.

In terms of this definition it would be possible to rate two StOs as similar even if they did not have any elements in common but

---

[5] To the author's knowledge, so far only the studies of Miller (1958), Garner and Whitman (1965), Whitman (1966), Koeppel (1967, 1968), and Runquist (1968a, 1970a) have taken letter position into account in verbal learning experiments when dealing with similarity.

only common spatial relationships among different elements. For example, two chairs which do not contain any identical or common elements would still be rated as similar because of the common spatial relationships among their different elements. The very fact that a common label, 'chair', is used to designate such StOs suggests that they are, in fact, similar.[6] The same considerations, of course, would apply to StOs which had only common sequential constraints (e.g. HHB and GGX), or only common elements (e.g.GYX and XIG). However, whereas common spatial relationships, either singly or in conjunction with common elements, lead to a relatively high degree of similarity (chair A - chair B, XIF-XIL), common elements by themselves, as was shown, do not lead to any significant degree of similarity (house-pile of prefabricated parts; GYX-XIG). Thus, although maximum similarity depends on the presence of all three parameters, each parameter does not contribute an equal share to similarity.

The use of such a similarity definition should lead to more consistent results between experiments using meaningful StOs and those using nonsense StOs. Nevertheless, one qualification has to be made. In verbal learning studies StOs appear in lists of more than two items in length. Therefore, two ranges of similarity must be distinguished: the similarity between any two StOs, and the similarity among all the StOs in the list. The new similarity definition deals only with the former range. However, it is the joint action of both that has to be

---

[6] Attributes such as common use (an environmental characteristic) do not have to be considered since other StOs are used for sitting without being called 'chairs'.

considered when the effect on learning is to be predicted.[7] In order to obtain the overall similarity of a list a simple pair-wise comparison of each item to every other item in the list cannot be used since it could yield a misleading estimate of overall similarity. For example, in some instances each pair-wise comparison could indicate a very high degree of similarity (VUK-VUL); however, the overall similarity could, in effect, be nil. List 1 in Figure 3 illustrates this case. When each item is compared with every other item the common VU in BOTH StOs makes them highly similar, particularly since the common letters are in common positions. On the other hand, when the list is considered as a totality the VU is COMMON to ALL StOs and therefore it is IRRELEVANT as far as the similarity of the whole list is concerned. In effect, the items in List 1 are all DIFFERENT. This is quite consistent with the usual evaluation of similarity among a number of items. For example, List 2 is considered as consisting of different items, and such COMMON factors as the following are regarded as obvious and therefore irrelevant to the total similarity in the list. All the items in List 2

(a) have a vowel in the middle,

(b) are written in upper case letters,

(c) are printed,

(d) are in black ink,

(e) are organized in horizontal lines,

(f) have three letters each,

etc.

---

[7] The formal similarity definition is applicable to both ranges. Nevertheless, the disadvantages far outweigh the advantages.

List 1	List 2	List 3
VUH	BOF	H
VUN	CEH	N
VUL	FUX	L
VUQ	NID	Q
VUG	MYP	G
VUC	VAW	C

Figure 3. Lists Varying in Similarity at Different Levels of Detail.

If characteristics that are common to ALL items in the set are to be disregarded as not contributing to the overall similarity of the list then in the same way the common VU-letters are also irrelevant. In effect, the list is reduced to H, N, L, Q, G, and C. All these items are different. Consequently, overall similarity cannot be obtained only by pair-wise comparisons. A variable is needed which would give low estimates of overall similarity in such cases and which would use pair-wise similarity in some way. The only available variable meeting these criteria is similarity based OP. Thus overall similarity may be defined in terms of the OP of a list using the pair-wise similarity as the basis for the organization. In List 1 there is no OP due to the common VU's, that is, the VU's cannot be used to organize the list. Therefore, overall similarity is nil. Thus if a list of *n* items is to have overall similarity, the pair-wise similarity must be such as to enable the formation of more than one group of *n* items and less than *n* groups of one item.

Overall similarity, defined in terms of the OP of a list, should have the same characteristics and effect on learning as does meaningful similarity. For example, a list consisting of musical instruments, animals, and professions is high in similarity, and permits clustering; a list consisting only of musical instruments in not high in similarity relative to the universe of the list. That all the items happen to be highly similar to each other when considered in relation to all the other items that might have been selected for the list is not pertinent to the estimate of the similarity among the items in the list. In other words, the overall similarity of a list depends on the existence of 'difference' in the list. Similarity cannot exist if difference

does not exist. Thus the list forms a microcosm and similarity in it depends on the difference within it. The common VU's are only part of the 'ground' of the list. They do not contribute to the similarity or to the difference in the list.

This becomes clear if the list consisting only of musical instruments is considered. In order to organize the list the Ss have to look for similarity among the DIFFERENT types of instruments: grouping the wind instruments into one set and the string instruments into another set.

Such grouping, however, is at a lower level of detail and it is this that has an effect on performance. In general, it may be hypothesized that performance will become increasingly inferior as the Ss have to organize on finer and finer details or at a lower level.

In summary, overall similarity is a relative concept depending for its degree (high, medium, and low) only on the universe in question which is the list. Therefore, its effect on learning will depend on the level of detail which has to be used in order to organize a given list. Consequently, the crucial factor in learning is the organization of the items in the list rather than the identity of the parts. Therefore, the effect of similarity on learning can best be understood in terms of the organizational potential that pair-wise similarity provides in the list at a given level of detail.

Chapter 6

THE EXPERIMENT

A learning experiment was designed to test four hypotheses. The first three hypotheses are Proposition I, II, and III, which were discussed in Chapter 4, and the fourth is the logical conclusion deduced from them. Although the former were amply supported with evidence from the literature, their importance as the theoretical foundation for the fourth hypothesis warranted that they also be tested experimentally.

Hypothesis I. Subjects do not primarily learn by rote, but seek organizational strategies whereby they can make the learning task easier.

Hypothesis II. Similarity has potential for organization.[1]

Hypothesis III. Organization of the material to be learned facilitates learning.

Hypothesis IV. Similarity facilitates learning under conditions of opportunity.[2]

---

[1]Hypothesis II is listed as a hypothesis even though it was not tested directly but only by inference. This treatment was selected because it was felt that it formed an integral part of the logical structure and therefore should not be separated from the other propositions.

[2]For a discussion of 'conditions of opportunity' see pp. 77-78, and pp. 80-92.

Definition of Variables

The experimental variables in this study are similarity, formal similarity, and organizational potential (OP).

1. Formal Similarity. This is defined in terms of the number of common letters per list. For a given number of items, similarity increases as the number of different letters decreases.

2. Similarity. There are two domains of similarity which have to be differentiated: the similarity between pairs of items, and the overall similarity of the list. This distinction has to be made since an increase in pair-wise similarity does not necessarily result in an increased degree of overall similarity.[3]

(a) Pair-wise similarity is defined in terms of the number of common elements in common positions with common sequential constraints, the latter referring to contiguous sequences only. The relationship between the three variables can be systematized in a three dimensional matrix. However, since the exact weighting of the dimensions is still to be determined, the definition remains qualitative. It can, however, be used if items are selected either along the diagonal or along any one of the dimensions, the other two being held constant.

One qualification has to be made however. The definition and the resulting matrix assume that the nonsense syllables are one dimensional stimulus objects (StOs). That is, all the items in a list are written horizontally on a plane. However, if the trigrams are defined as three dimensional StOs and rotations out of the plane of the paper

---

[3]See p.114 for a discussion of this point.

are permitted, then mirror-images such as VUK-KUV become highly similar which they are not when regarded strictly in a one dimensional way.

(b) Overall list similarity is defined in terms of the degree of OP using pair-wise similarity as a basis. Consequently, overall list similarity increases with increased pair-wise similarity but only as long as OP increases as well. Therefore, a list such as #1, Figure 3, p. 116, although high in pair-wise similarity would be low in overall similarity since the pair-wise similarity cannot be used for organization, that is, OP is low.

(3) Organizational potential (OP). This variable is designated as a 'potential' since it does not guarantee that the available structure will be perceived or used.

Organizational potential is defined in terms of groupability, comprehensiveness, and complexity.

(a) Groupability means the possibility of grouping the items in the list into more than one group of all *n* terms, and into less than *n* groups of one item. Groupability, however, does not indicate the basis for the formation of the groups nor the relationship, if any, between the groups. Thus, items may be grouped on the basis of similarity; or, as in List 3, Figure 4, on the basis of alphabetical proximity (B and C; P and Q; and M and N). Lists, 1, 2, and 4 illustrate grouping on the basis of similarity.

(b) Comprehensiveness is defined in terms of the degree to which all the items in the list can be encompassed in one simple coding statement. Thus comprehensiveness is enhanced if a relationship exists between the groups so that they can be structured into a bigger

List 1

VUK
VUX
VUT

ZOH
ZOS
ZOJ

List 2

VUK
VUX
VUT

HOZ
SOZ
JOZ

List 3

BOF
CYZ

MYP
NID

POG
QIJ

List 4

VUK
VUL
VUT

QIF
QAF

GYX
MYP

HOZ
SOZ
JOZ

List 5

VUX
VUZ

XIL
XIF

ZOS
ZOH

List 6

VUX
VUZ

XIL
XIZ

ZOS
ZOV

Figure 4. Lists Varying on the Defining Characteristics of Organizational Potential (OP).

unit, in other words, so that the 'chunks' can be recoded into bigger 'chunks'. Therefore, List 1, Figure 4, has the most comprehensiveness since not only do both groups contain an equal number of items but the locus of similarity is the same in both groups. That is, the common letters occur in the first and second position in both groups. List 2, is a little lower in comprehensiveness since an added qualification is needed to identify the locus of similarity. List 4 is the lowest in comprehensiveness as the items in it can neither be grouped into equal item groups nor can they be related with a simple statement into one big group. List 3 is like List 1 but not entirely so. Although a common rule may be used to characterize the relationship within each group, the alphabetical nature of the grouping is not maximized. Since the basis of organization is a learned sequence the comprehensiveness of the list would have been enhanced if the first member in each pair had been an equal distance from the first member in the other two groups, that is, if the interval between B and M had been the same as that between M and P on the alphabetical scale.

(c) <u>Complexity</u> refers to the degree and quality of the multiplicity of available alternatives for the formation of primary and secondary groups. In other words, if the items in the list can be members of more than one group, complexity is increased and OP may be decreased depending on the quality of the available categories. If by belonging to more than one group an item can contribute to the regularity of the overall structure then OP will be enhanced. For example, in List 4, Figure 4, the primary grouping is on the basis of first and second letter identity; however, VUX could be a member of group 2 on the basis of common letters while VUZ could also be a member of group 3

because of the common Z. Nevertheless, the quality of the complexity is regular and not random since both letters of the first group are used as the initial letters of the other two groups. Had the complexity been random, as in List 6, then OP would have been reduced because of the increase in the complexity of the coding system necessary to specify the structure.

Consequently, high organizational potential (HOP) is characterized by the simplicity of the final coding statement.

As it stands the definition is qualitative rather than quantitative; however, there should be no difficulty in applying it for gross categorizations such as high, moderate, and low OP.

The Materials

A questionnaire and eight experimental lists were constructed. The purpose of the questionnaire was to determine the learning strategies, if any, employed by the subjects (*Ss*) during the first part of the experiment. In addition it was to provide some indication as to the extent to which the *Ss* coded the StOs as pronounceable units rather than adopting the alternative of spelling them. Finally, it was to give the *Ss*' reaction to the sequence in which the items were presented in the list. A specimen copy of the questionnaire appears in Appendix F.

The learning materials consisted of two sets of lists which differed only in the sequential ordering of the items. One set was designated as the ordered set (S) and the other as the random set (R). Each set consisted of four lists of 12 CVC trigrams each, matched on frequency and association values, the latter ranging from 0 to 47 on

the Glaze scale (Underwood and Schulz, 1960).

The lists were constructed so as to vary with respect to formal similarity, similarity, and OP. Table 3 gives the list descriptions in terms of the three experimental variables. The inverse relationship between similarity and OP at the low end of the spectrum occurs as a result of other bases of organization present in the 400 list. In other words, the OP value given in the table is _general_ and not just similarity based. The column designated 'similarity' indicates the joint effect of pair-wise similarity and _similarity based_ OP.

The degree of pair-wise similarity for each list is given in Table 4 and all the lists are reproduced in Table 5. A specimen set of all the lists also appears in Appendix E.

The lists were constructed according to the following specifications:

List 100. This is the high similarity (HS), high formal similarity (HfS), and high organizational potential (HOP) list. It was constructed from ten letters: eight consonants and two vowels. Essentially the list consists of two StOs and their 180° rotations or mirror-images. Consequently, the degree of pair-wise similarity is very high (Table 4). In terms of OP the description can be coded as follows:

$$\left.\begin{array}{ll} \text{(VU)} & \text{(K, T, X)} \\ \text{(ZO)} & \text{(H, J, S)} \end{array}\right\} \text{ and the six mirror-images.}$$

Thus, similarity based OP is also rather high.

TABLE 3

DESCRIPTION OF EXPERIMENTAL LISTS IN TERMS OF FORMAL SIMILARITY, SIMILARITY AND ORGANIZATIONAL POTENTIAL

Name of List	Degree of Formal Similarity (fS)	Degree of Similarity (S)	Degree of Organizational Potential (OP)
100 (HSL)	high	high	high
200 (MSL)	high	moderate	moderate
300 (LSL)	high	low	very low
400 (NSL)	almost nil	almost nil	low

TABLE 4

DESCRIPTION OF PAIR-WISE SIMILARITY WITH FREQUENCY OF OCCURRENCE PER LIST

Category[a]	Frequency of Pairs Per List 100	200	300	400	Description of Similarity
1	6	6			Mirror-image pairs
2	12	6		1	(2,2,2)-2 common elements in 2 common positions with common sequential constraints
3	12	6	4		(2, 1, 0)
4			19	6	(1, 1, 0)
5			1		(2, 0, 0)
6			8		(1, 0, 0)
Sub-Total:	30	18	32	7	
7	36	48	34	59	(0, 0, 0)
Total:	66	66	66	66	

[a]The categories are listed in decreasing degrees of similarity.

TABLE 5

EXPERIMENTAL LISTS

100 S	200 S	300 S	400 S
VUK	QIF	FUX	BOF
VUT	QIJ	HUZ	CEH
VUX	FIQ	JIH	GYX
KUV	JIQ	KUJ	LAJ
TUV	VUK	QIJ	MYP
XUV	VUX	QOS	NID
ZOH	KUV	SOZ	QIJ
ZOJ	XUV	VOF	QUR
ZOS	ZOH	VUK	TEH
HOZ	ZOS	XIF	VAW
JOZ	HOZ	XOQ	VUK
SOZ	SOZ	ZIV	ZOS

100 R	200 R	300 R	400 R
JOZ	HOZ	XOQ	VUK
ZOH	KUV	SOZ	QIJ
HOZ	ZOS	XIF	VAW
VUK	QIF	FUX	BOF
SOZ	SOZ	ZIV	ZOS
TUV	VUK	QIJ	MYP
VUT	QIJ	HUZ	CEH
XUV	VUX	QOS	NID
ZOJ	XUV	VOF	QUR
KUV	JIQ	KUJ	LAJ
ZOS	ZOH	VUK	TEH
VUX	FIQ	JIH	GYX

List 200. This is the moderate similarity (MS), high formal similarity (HfS), and moderate organizational potential (MOP) list. It was constructed from 12 letters: nine consonants and three vowels. Essentially the list consists of three StOs and their mirror-images. In terms of the OP description this can be summarized as follows:

$$\left.\begin{array}{ll} (QI) & (F, J) \\ (VU) & (K, X) \\ (ZO) & (H, S) \end{array}\right\} \text{and the six mirror-images.}$$

Thus, similarity based OP is somewhat less than for List 100. The degree of pair-wise similarity is also less (Table 4). Therefore, the overall similarity of List 200 is somewhat less than that of List 100. The degree of difference is not large; however, the requirements of matching necessitated this construction.

List 300. This is the low similarity (LS), high formal similarity (HfS), and very low organizational potential list (VLOP). It was constructed from 12 letters: nine consonants and three vowels. Table 4, page 127, gives its pair-wise similarity. In terms of number of pairs that were totally different, List 300 is matched with List 100. The OP of this list cannot be described in any simple way and thus will not be attempted.

List 400. This is the no similarity (NS), no formal similarity (NfS) and low organizational potential (LOP) list. The list is designated as 'no' even though there is some similarity. This was unavoidable with a 12 item list. Thus, all the letters of the alphabet were used with H, J, Q, and V being used twice. The letter Y was used

only as a vowel. All the vowels were used twice. The OP is designated as low rather than nil because there was some similarity which could be used as a basis for organization, and the contiguous alphabetical sequence that some of the items formed also provided some additional OP. For example, LAJ, MYP, and NID were in alphabetical order. The list, however, could not be reduced to one simple group. At best there were four groupings available. These may be represented as follows:[4]

```
                    ZOS
                  . BOF
                  . CEH - TEH
         . LAJ
         . MYP - GYK      QIJ - QUR      VAW - VUK
         . NID
```

This illustrates the use of a hybrid OP, combining similarity with alphabetical order. Such groupings have, however, a disjunctive character since group membership is described in terms of more than one property. For example, the top group is formed if an item has an 'O' or an 'EH' or is in alphabetical sequence (B and C). Such disjunctive groupings, however, are not formed as readily as conjunctive groups (Bruner, Goodnow, and Austin, 1956).

In general, OP is low for the 400 list but not as low as for the 300 list since in the latter case it is not possible to describe the OP in a simple way because many items can belong to more than one group and the relationship is not consistent but is random.

---

[4]This is a summary of the most common strategies that were reported by the <u>Ss</u> on the questionnaire.

The Subjects

The Ss were 299 senior high school students attending Leduc Senior High School, Leduc; St. Albert High School, St. Albert; and Memorial Composite High School, Stony Plain, Alberta. All Ss were naive with respect to verbal learning experiments. A one-way analysis of variance indicated that Schools was not a significant variable at the 0.10 level on all 23 performance measures used in the experiment.[5] Therefore, the final sample of 240 Ss was selected from the total group of 287 Ss, 12 Ss having been rejected for not following instructions. The method of selection was random except for the following conditions:

(a) An approximately equal number of Ss were selected from School 2 (St. Albert) and from the combined pool of Ss at School 1 and 3 (Leduc and Stony Plain respectively).[6]

(b) Exactly 15 Ss of each sex were selected for each treatment combination.

Procedure

The experiment was conducted during regular school periods with eight groups of Ss ranging in size from 22 to 67.

The supervisor began the session by giving the Ss some background

---

[5]For three performance measures the homogeneity of variance assumption was not met. However, the calculation of the S-statistic indicated that the critical F-value was not affected to any great extent and the difference between Schools for the three measures was not significant at the 0.05 level. A discussion of the S-statistic is given in Appendix G, and the statistical data are summarized in Appendix H.

[6]Condition (a) was imposed because the variance on a number of measures for School 2 was smaller than that for the other two Schools, and in three cases, as indicated in the previous footnote, the difference between the variances was significant at the 0.05 level.

information as to the nature of the learning experiment (Appendix B). Next each S received the following:

(a) A Practice Booklet (Appendix D)

(b) A Questionnaire in a brown envelope (Appendix F)

(c) An Instruction sheet (Appendix C)

(d) A Word List, placed face down on the desk (Appendix E)

One set of Word Lists was handed out randomly to the female Ss and one set to the male Ss. In this way an approximately even distribution of the sexes in each treatment combination was achieved.

After the supervisor read the instructions to the class, the Ss practiced the learning method using the four three-digit numbers provided in the example. When all questions as to procedure were answered, the actual learning experiment began and continued until all Ss had reached the criterion of one perfect recall on trial 5, 10, or 15.

The Ss were given 2.0 minutes to study the Word List and 1.0 minute to write down in the Practice Booklet as many of the nonsense words as they could remember. At the end of the fifth trial (a study-recall sequence being considered as one trial) the Ss were asked to mark their latest recall. If all 12 items were correct the S proceeded to the Questionnaire, if not, then the S continued for five more trials and the same procedure was followed. At the end of the 15 trials all Ss went on to the Questionnaire even if they had not reached criterion.

After completion of the Questionnaire, the Ss placed all the materials into the envelopes which had been provided with the Questionnaire. All envelopes were then collected by the supervisor.

Scoring Recall and Questionnaire

In scoring the recall protocols 24 performance measures were

recorded per S (Table 6).

Although the Ss went up to 5, 10, or 15 trials, all intervening trials were marked up to the first correct trial, except that the first five trials for all Ss were marked. The 'Trials to Criterion' was the number of the trial on which the S had all items correct for the first time.

Since recall was free the sequence of items was not a determining factor of correctness.

Three types of errors were distinguished and scored separately:

(a) Overt Errors

(b) Omissions

(c) Unclassifiable (incomplete trigrams and duplicates)

Category (a) and (b) occured because the Ss were not forced to write 12 responses on each trial. Consequently, two types of errors were possible, those of commission and those of omission.

The scoring of the questionnaire was done in four parts:

(a) Learning Strategy, questions 1, 3, 4, 6, and 7

(b) Similarity Based Associations, question 5

(c) Pronounceability of Items, question 8

(d) Preferred Order, question 9

An examination of the verbal reports dealing with learning strategy suggested seven different categories with a number of sub-classifications. The categories with their sub-classes are presented in Table 7. A code is used to differentiate between 'comprehensive' and 'limited' strategies. The former designation indicates those strategies which were used to organize all 12 items in the list, for example, the grouping of the items in a list into three groups of four, or the use

TABLE 6

PERFORMANCE MEASURES[a]

1. Number of Trials to Criterion

2. Number Correct on Trial #1
3. Number Correct on Trial #2
4. Number Correct on Trial #3
5. Number Correct on Trial #4
6. Number Correct on Trial #5

7. Total Number Correct on Five Trials (#1-#5)

8. Number Wrong (Overt Errors) on Trial #1
9. Number Wrong (Overt Errors) on Trial #2
10. Number Wrong (Overt Errors) on Trial #3
11. Number Wrong (Overt Errors) on Trial #4
12. Number Wrong (Overt Errors) on Trial #5

13. Total Number Wrong (Overt Errors) on Five Trials (#1-#5)

14. Number Wrong (Omissions) on Trial #1
15. Number Wrong (Omissions) on Trial #2
16. Number Wrong (Omissions) on Trial #3
17. Number Wrong (Omissions) on Trial #4
18. Number Wrong (Omissions) on Trial #5

19. Total Number Wrong (Omissions) on Five Trials (#1-#5)

20. Mean Number Correct per Trial

21. Mean Number Wrong (Overt Errors) per Trial

22. Mean Number Wrong (Omissions) per Trial

23. Mean Number Wrong (Unclassifiable) per Trial

24. Mean Total Errors per Trial

[a]All are means.

TABLE 7

A DESCRIPTION OF THE LEARNING STRATEGY CLASSIFICATION SYSTEM

Strategy	Use[a]		Description of Sub-Classification
1. Similarity	C	1.1	Grouping into pairs of mirror-images, e.g. VUK - KUV
	C	1.2	Grouping into pairs or threes by first two or last two letters, e.g. QIF QIJ or VUK VUT VUX or JOZ HOZ SOZ
	C	1.3	Grouping into fours or sixes by common letters and mirror-images, e.g. FIQ - QIF JIQ - QIJ or KUV - VUK TUV - VUT XUV - VUX
	C	1.4	Grouping six 'basic' words together (the other six being the mirror-images to the 'basic' words)
	CS	1.5	Use of letter relationships, or generative strategy
	L	1.6	Grouping by common letter in common position, (list 300 & 400 only), 300 list: only 3 mutually exclusive and 2 overlapping pairs available: VOF VUK & XIF XOQ & HUZ SOZ & KUJ QIJ & QIJ QOS (note QIJ)

(continued)

[a]The code indicates the extent to which strategy was applied, or any special considerations.

Key: C -(comprehensive strategy)-used to organize all 12 items
L -(limited strategy)-used to organize a few items only
S -(special case)-sub-class is not a proper member of the class

TABLE 7 (continued)

Strategy	Use		Description of Sub-Classification
1. Similarity	L	1.6	(continued) OR grouping into three groups of four by middle vowel[b] 400 list: only 2 mutually exclusive and 2 overlapping pairs available: CEH & VAW & LAJ & QIJ TEH VUK QIJ QUR (note QIJ) AND also available: 2 pairs of mutually exclusive vowels (third is with CEH-TEH) and three overlapping pairs[c]: e.g. BOF & GYX & NID & QIJ ZOS MYP QIJ QUR etc.
	L	1.7	Grouping by common letters in mirror positions (only in 300 list), 300 list: only 3 pairs of two letter reversals and 1 pair of first-last reversal, VUK & QIJ & XOQ & XIF KUJ JIH QOS FUX
	L	1.8	Grouping by #6 AND #7 (list 300 only)
	S	1.9	Similarity noticed only, no indication given as to use OR similarity noticed but NOT used
2. Aural Grouping	L	2.1	Grouping into similarly sounding pairs e.g. KUJ - QIJ

[b]Although three groups of four would be a comprehensive strategy it is included under limited strategies because it was only used for one group of four items.

[c]Although 6 pairs would be a comprehensive strategy, vowel grouping was used for less than all pairs, therefore, it is designated as 'Limited'.

TABLE 7 (continued)

Strategy	Use		Description of Sub-Classification
3. Chaining	L	3.1	Forming a chain of 2, 3, or 4 items by linking last and first letter e.g. SOZ - ZIV - VUK - KUJ
4. Numerical Grouping	C	4.1	Grouped into 2 groups of six
	C	4.2	Grouped into 3 groups of four
	C	4.3	Grouped into 4 groups of three
	C	4.4	Grouped into 6 groups of two
	C	4.5	Grouped into two or three odd groups e.g. 8 and 4 3, 4, and 5
5. Rhyming or Use of Sound	L	5.1	Grouping into rhyming pairs e.g. M - Y - P, N - I - D or TEH and CEH
	L	5.2	Use of a 'sound pattern' or 'sound in a series' etc.
	L	5.3	Miscellaneous e.g. sounded funny
6. Alphabetical Order	C	6.1	All 12 items in alphabetical order
	L	6.2	Some items in alphabetical order (more than four)
	L	6.3	One pair in alphabetical order, that is, two items grouped together because they were in alphabetical sequence

(continued)

TABLE 7 (continued)

Strategy	Use		Description of Sub-Classification
6. Alphabetical Order	L	6.4	Individual items recalled because of association with alphabet, e.g. JIH is in reverse order H,I,J, or TUV is in order T,U,V.
7. Miscellaneous	L	7.1	Use of sentences for a few items or for individual letters in one item
	C	7.2	Grouping by difficulty
	C	7.3	Grouping by pronounceability
	C	7.4	Differentiated practice
	C	7.5	Grouping <u>letters</u> together by position (first column, second column, and third column of letters)
	C	7.6	Visualizing total list
	C	7.7	Various sequences and orders e.g. Learning from last to first or learning from middle up and down
	L	7.8	Miscellaneous personal associations and mnemonic devices
	L	7.9	Others

of alphabetical order for all 12 items. The latter designation refers to those strategies which were used for seven items at the most, for example, the formation of a four item chain such as SOZ-ZIV-VUK-KUJ, or the use of a rhyme pattern between two CVC's when they are spelled as in the case of V-U-K, Q-I-J. In some instances the 'limited' designation indicates the use of a strategy for a single item, for example, the use of alphabetical order for TUV (T, U, V).

Questions 5, 8, and 9 were scored in a straight-forward manner by recording the _Ss'_ answers directly.

Chapter 7

RESULTS

The primary data arose from four performance measures on four lists by two item orders. The lists varied in formal similarity, similarity, and organizational potential (OP) (see Table 3, p. 126). The items in each list were either organized (S) or in a random sequence (R).

The secondary data arose from the Questionnaire which was administered after the learning-recall phase was completed. The purpose of the Questionnaire was to determine the Ss' learning strategy, their reaction to the item sequence in the lists, and their predominant coding mode (spelling versus pronunciation of CVC's).

Analysis of Primary Data

The four performance measures were the number of trials to criterion, the mean number correct per trial, the mean number of overt errors per trial, and the mean number of omission errors per trial. The data are summarized in Tables 8, 10, 12, and 14. A two-way analysis of variance indicated that the list variable was highly significant on all performance measures ($p < 0.0001$). The order variable, however, was only significant for trials to criterion ($p < 0.0002$), mean number correct per trial ($p < 0.03$), and mean number of omission errors per trial ($p < 0.002$). Order did not reach significance at the 0.05 level for the mean number of overt errors. The list by order interaction was not

TABLE 8

MEAN NUMBER OF TRIALS TO CRITERION FOR
ORDERED (S) AND RANDOM (R) LISTS
COVARYING IN SIMILARITY AND
FORMAL SIMILARITY

List	Mean Number of Trials to Criterion	
	(S)	(R)
100[a]	2.47	3.73
200	2.30	4.33
300	7.00	7.77
400	3.53	4.53

[a]See pp. 125-129 for key.

TABLE 9

SUMMARY OF TWO-WAY ANALYSIS OF VARIANCE
FOR TRIALS TO CRITERION

Source of Variation	SS	df	MS	F
A (Lists)	713.09	3	237.70	36.31*****
B (Order)	96.27	1	96.27	14.71***
AB	13.63	3	4.54	0.69
Error	1518.61	232	6.55	

*** Significant at 0.001 level, $p < 0.0002$

***** Significant at 0.00001 level, $p < 0.000002$

TABLE 10

MEAN NUMBER CORRECT PER TRIAL FOR ORDERED (S) AND RANDOM (R) LISTS COVARYING IN SIMILARITY AND FORMAL SIMILARITY

List	Mean Number Correct Per Trial	
	(S)	(R)
100[a]	10.26	10.25
200	10.52	9.36
300	8.26	8.12
400	9.78	9.53

[a]See pp. 125-129 for key.

TABLE 11

SUMMARY OF TWO-WAY ANALYSIS OF VARIANCE FOR MEAN NUMBER CORRECT PER TRIAL

Source of Variation	SS	df	MS	F
A (Lists)	150.09	3	50.03	26.77*****
B (Order)	9.05	1	9.05	4.84*
AB	12.15	3	4.05	2.17
Error	433.59	232	1.87	

* Significant at 0.05 level, $p < 0.03$

***** Significant at 0.00001 level, $p < 0.000003$

TABLE 12

MEAN NUMBER WRONG (OVERT ERRORS) PER TRIAL FOR ORDERED (S) AND RANDOM (R) LISTS COVARYING IN SIMILARITY AND FORMAL SIMILARITY

List	Mean Number Wrong (Overt Errors) Per Trial	
	(S)	(R)
100[a]	0.73	0.52
200	0.87	0.85
300	1.23	1.27
400	0.71	0.72

[a]See pp. 125-129 for key.

TABLE 13

SUMMARY OF TWO-WAY ANALYSIS OF VARIANCE FOR MEAN NUMBER WRONG (OVERT ERRORS) PER TRIAL

Source of Variation	SS	df	MS	F
A (Lists)	13.71	3	4.57	8.26****
B (Order)	0.10	1	0.10	0.19
AB	0.57	3	0.19	0.35
Error	128.37	232	0.55	

**** Significant at 0.0001 level, $p < 0.00004$

TABLE 14

MEAN NUMBER WRONG (OMISSIONS) PER TRIAL
FOR ORDERED (S) AND RANDOM (R) LISTS
COVARYING IN SIMILARITY AND
FORMAL SIMILARITY

List	Mean Number Wrong (Omissions) Per Trial	
	(S)	(R)
100[a]	0.86	1.05
200	0.54	1.64
300	2.26	2.54
400	1.44	1.67

[a]See pp. 125-129 for key.

TABLE 15

SUMMARY OF TWO-WAY ANALYSIS OF VARIANCE
FOR MEAN NUMBER WRONG (OMISSIONS)
PER TRIAL

Source of Variation	SS	df	MS	F
A (List)	76.39	3	25.46	23.23*****
B (Order)	12.02	1	12.02	10.96**
AB	8.41	3	2.80	2.56
Error	254.29	232	1.10	

** Significant at 0.01 level, $p < 0.002$

***** Significant at 0.00001 level, $p < 0.000004$

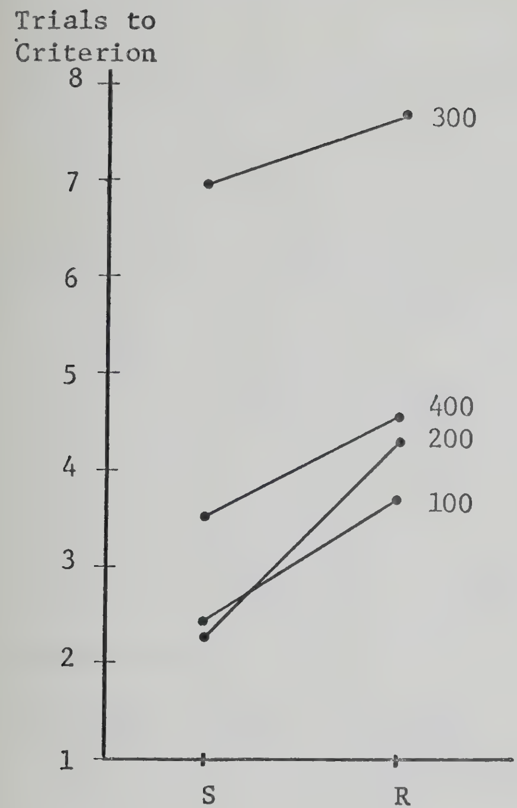
significant at the 0.05 level ($F_{.95} = 2.64$) for all four performance measures. However, for the mean number of omission errors per trial and for the mean number correct per trial the interaction did approach significance ($F = 2.56$, $p < 0.06$ and $F = 2.17$, $p < 0.10$ respectively). The statistical analyses are summarized in Tables 9, 11, 13, and 15.[1]

For each performance measure the difference between the means for the two levels of the order variable (S and R) was analysed at each level of the list variable (100, 200, 300, 400) using four orthogonal comparisons.[2] The results of the analysis indicated that for all performance measures, except for the mean number of overt errors per trial, the difference between the S and R item order was significant only for the 200 list. An examination of Figure 5a shows that list 200 is the one which is affected most by a random arrangement of the items. Under the S condition 2.30 trials were needed to reach criterion; however, under the R condition 4.33 trials were required. This is almost twice as many. The difference between the means is significant at the 0.01 level ($t = 3.08$, $t_{.99} = 2.60$). On the other hand, list 100 is not affected as much: 2.47 trials for the S condition and 3.73 for the R condition. The difference does nor quite reach significance at the 0.05 level ($t = 1.92$, $t_{.95} = 1.97$). The two lists differ in terms of OP and similarity, list 100 being highest in both. Thus, the higher the similarity and OP the less the negative effect of randomization.

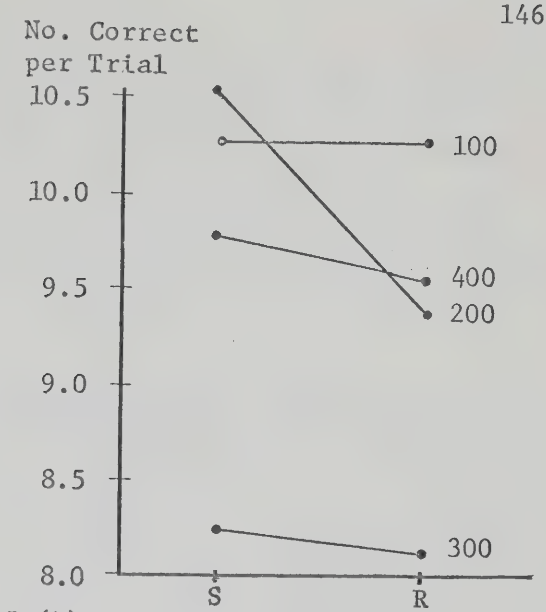
---

[1]In all analyses the denominator for the F-ratio was calculated on the basis of a fixed effect model with equal n's per cell (ANOV25, computer program).

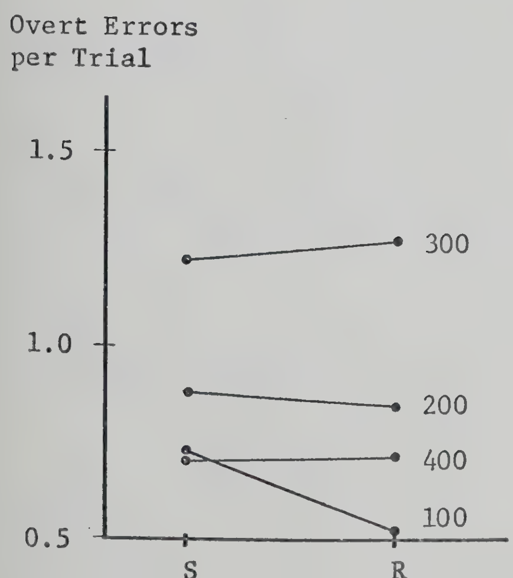
[2]For a discussion of planned orthogonal comarisons see Edwards (1960) pp. 140-148 and pp. 189-191. A sample calculation and a summary of the statistical analysis is given in Appendix K.



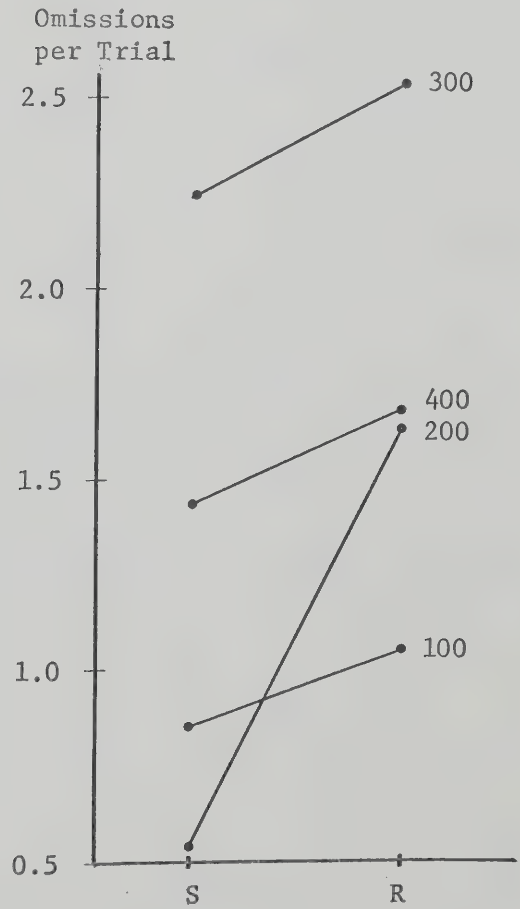
5 (a). Profile of Data (Table 8).



5 (b). Profile of Data from Table 10.



5 (c). Profile of Data (Table 12).



5 (d). Profile of Data from Table 14.

Figure 5. Profiles for the Order Variable at Four Levels of Similarity for (a) Trials to Criterion, (b) Mean Number Correct per Trial, (c) Mean Number of Overt Errors, and (d) Mean Number of Omission Errors per Trial.

On the other hand, list 300 which is low in similarity and very low in OP was affected least. The mean number of trials to criterion were 7.00 and 7.77 for the S and R conditions respectively ($t = 1.16$). Consequently, randomization has the least effect at both extremes of the similarity and OP continuum.

List 400 which was the no-similarity list and low in OP (but higher than the 300 list) was affected somewhat more than the 300 list. For the S condition 3.53 trials were needed and for the R condition 4.53 trials ($t = 1.51$). On the average, one extra trial was needed to overcome the effects of randomization. Thus the relationship between OP and order is quite consistent. At very high OP randomization has some effect, at intermediate OP the hindering effect reaches a maximum, and as OP decreases to low levels the effect also decreases and then almost disappeares. A similar pattern was observed for the other two performance measures for which the overall effect was significant. For the mean number correct per trial (Figure 5b) the difference between the S and R condition for list 200 was significant at the 0.01 level ($t = 3.27$, $t_{.99} = 1.97$). The obtained t values for all other differences were less than one (see Appendix K). For the mean number of omission errors per trial (Figure 5d) the difference between the S and R condition for list 200 was significant at the 0.001 level ($t = 4.04$, $t_{.999} = 3.36$).

The relationship between the effects of randomization and similarity also follows the same pattern since similarity and OP (similarity based OP) vary directly. However, for the 400 list the effect is confounded because of non-similarity based OP.

The nature of the relationship between formal similarity and randomization is not systematic. Lists 100, 200, and 300 were made

equally high in formal similarity, but the effect of randomization was different for each.

In general the data indicate that a random arrangement of the items in the list leads to inferior performance, requiring more trials to reach criterion and resulting in fewer number correct per trial. However, whereas the number of omission errors per trial is significantly higher for the random lists, the number of overt errors is almost identical for both item orders. The relationship between the effects of randomization and OP (and similarity) was seen to be curvilinear, while that between formal similarity and randomization was found to be irregular.

The variation among the lists was explored further by means of the Scheffe Multiple Comparison technique.

Trials to Criterion. Pair-wise comparisons between each mean and every other mean indicated that only the differences between the 300 list and the 100, 200, and 400 lists were significant ($p<0.00001$). The means for the 100, 200, 300, and 400 lists were 3.10, 3.32, 7.38, and 4.03 trials respectively. As can be seen the 300 list required the most number of trials to reach criterion. Although the difference is not statistically significant the 400 list required the next fewer number of trials while lists 100 and 200 required the least number of trials. The performance parallels that of OP rather than that of similarity. The relationship between performance and formal similarity is not consistent, being at once the best and the poorest. A comparison of the mean number of trials for the 100 list with the mean number of trials for the 300 list shows this clearly. The best performance occurred on the 100 list (3.10 trials to criterion) while the poorest

performance occurred on the 300 list (7.38 trials to criterion); yet both lists were equal in formal similarity.

Mean Number Correct per Trial. The relationship was exactly the same as for trials to criterion. For lists 100, 200, 300, and 400 the number correct were 10.26, 9.94, 8.19, and 9.65 respectively. Again only the 300 mean was significantly different from each one of the others ($p<0.00001$).

Mean Number Wrong (Overt Errors) per Trial. The means were as follows: 0.63, 0.86, 1.25, and 0.71 for the 100, 200, 300, and 400 lists respectively. Again the same pattern is repeated; however, the differences are not as highly significant. The difference between 100 and 300, and between 400 and 300 is significant at the 0.01 level ($p<0.0002$, $p<0.002$ respectively). However, the difference between 200 and 300 barely reaches significance at the 0.05 level ($p<0.045$).

Mean Number Wrong (Omission Errors). More errors of omission were made than of commission, the means being 0.96, 1.09, 2.40, and 1.55 for the 100, 200, 300, and 400 lists respectively. Again the difference between 100 and 300, 200 and 300, and 400 and 300 were highly significant ($p<0.00001$, $p<0.00001$, $p<0.001$ respectively). However, the difference between 100 and 400 was also significant, $p<0.03$.

Summary. The differences between the means for the four levels of the similarity variable were not large enough to reach significance; however, the means do go in the expected direction, that is, performance tends to improve with increased similarity and/or OP. The lack of significant differences may be attributed to the two minute learning time interval per trial which was too long to permit for finer differentiations in performance. As far as formal similarity is concerned

the results were reasonably conclusive in indicating that the variable is not related in any consistent way with performance.

A subsidiary examination was carried out on the first five learning trials so as to determine the nature of the performance in the initial stages and to relate this to the variables of interest. The mean number correct, the mean number of overt errors, and the mean number of omission errors for each trial are given in Tables 16, 17, and 18 respectively.

The mean number correct as a function of trials is plotted in Figure 6. The relationship between lists and order becomes very clear after the second trial. On all five trials, lists and order are significant at the 0.05 level. Figures 7 and 8 show the relationship for the S lists and R lists separately, while Figures 9, 10 and 11 compare each of the lists with the 300 list. An examination of Figures 9, 10, and 11 indicates the consistent superiority of the ordered list on all five trials. Furthermore, Figure 10 shows the greater impact of randomization on the 200 list, and the somewhat lesser effects on the 100 and 400 lists (Figures 9 and 11). In comparison the effect on the 300 list is small. In terms of formal similarity lists 100, 200, and 300 were equally high; nevertheless consistent differences indicate that for high formal similarity, performance can be superior to the 400 list or inferior to it (Figures 7 and 8).

The relationship between high OP and moderate OP did not come out clearly and in fact the 200S list (moderate OP, moderate similarity) was consistently superior to the 100S list (high OP, high similarity), while in the R condition the reverse was true. Consequently, the overall difference between the 200 and 100 lists is not significant on any trial at the 0.05 level. In order to determine the cause of this

TABLE 16

THE (MEAN) NUMBER CORRECT IN EACH LIST x ORDER CONDITION FOR THE FIRST FIVE TRIALS

List & Order	Trials 1	2	3	4	5
100S	8.90	10.67	11.03	11.50	11.97
200S	9.17	10.73	11.60	11.97	11.97
300S	5.10	6.33	7.80	8.53	9.30
400S	7.23	9.90	10.97	11.37	11.77
100R	8.73	9.67	10.50	11.00	11.33
200R	7.30	8.60	9.90	10.60	11.10
300R	4.53	5.93	7.37	8.27	9.13
400R	6.93	8.90	10.30	10.97	11.47

TABLE 17

THE (MEAN) NUMBER WRONG (OVERT ERRORS) IN EACH LIST x ORDER CONDITION FOR THE FIRST FIVE TRIALS

List & Order	Trials 1	2	3	4	5
100S	1.27	0.53	0.33	0.33	0.03
200S	1.53	0.87	0.27	0.03	0.03
300S	2.23	2.03	1.00	1.07	0.97
400S	1.23	0.70	0.37	0.40	0.20
100R	0.53	0.90	0.53	0.37	0.40
200R	1.30	0.93	0.87	0.43	0.40
300R	1.90	1.33	1.40	1.57	1.30
400R	0.97	1.10	0.67	0.53	0.23

TABLE 18

THE (MEAN) NUMBER WRONG (ERRORS OF OMISSION) IN EACH LIST x ORDER CONDITION FOR THE FIRST FIVE TRIALS

	Trials				
List & Order	1	2	3	4	5
100S	1.57	0.80	0.57	0.10	0.00
200S	1.20	0.33	0.07	0.00	0.00
300S	4.63	3.50	2.73	2.13	1.63
400S	3.47	1.27	0.57	0.20	0.03
100R	2.50	1.40	0.83	0.43	0.17
200R	3.37	2.27	1.10	0.80	0.30
300R	5.47	4.70	3.10	2.10	1.53
400R	3.93	1.93	1.00	0.50	0.23

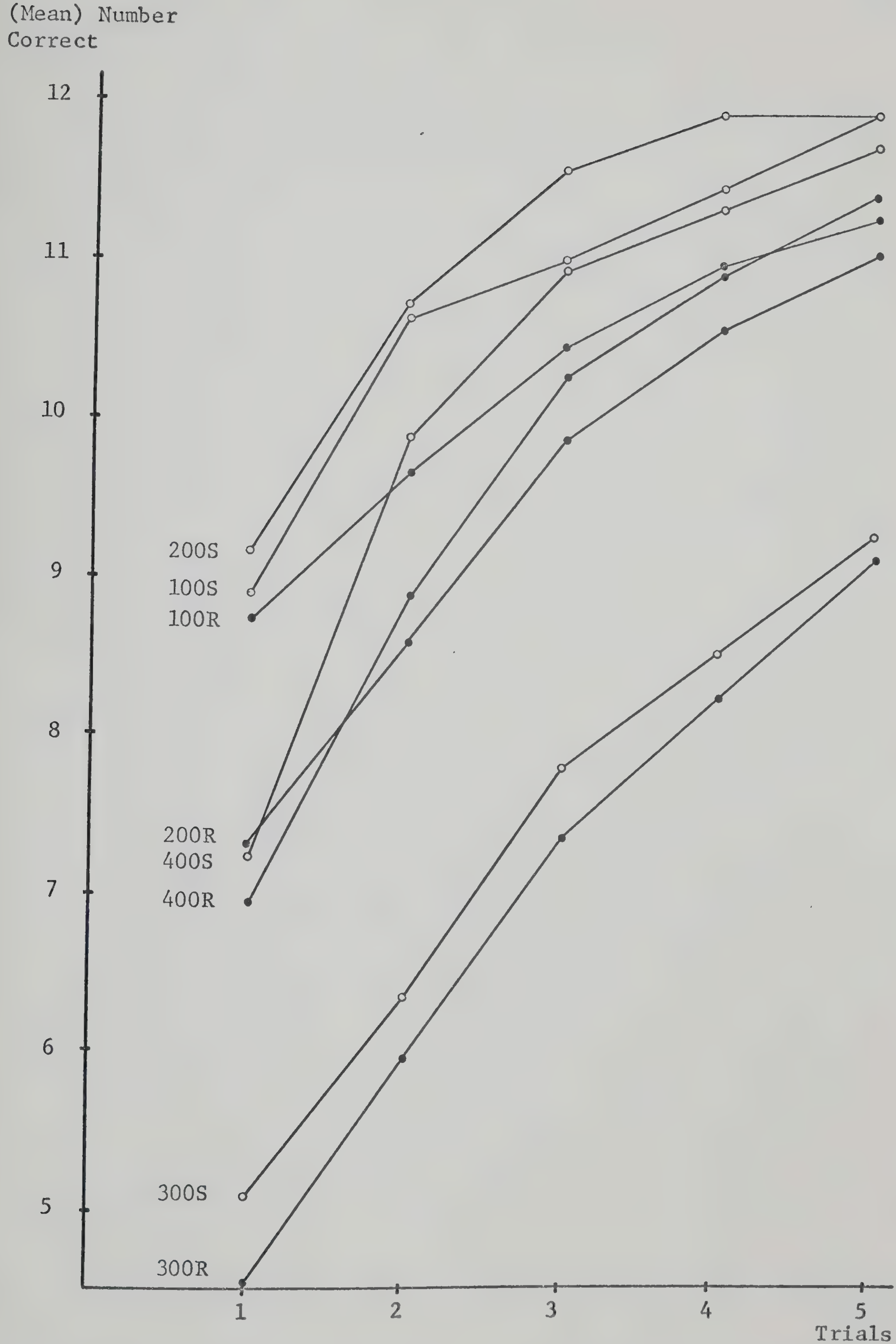


Figure 6. The (Mean) Number Correct on the First Five Trials for Random (R) and Ordered (S) Lists.

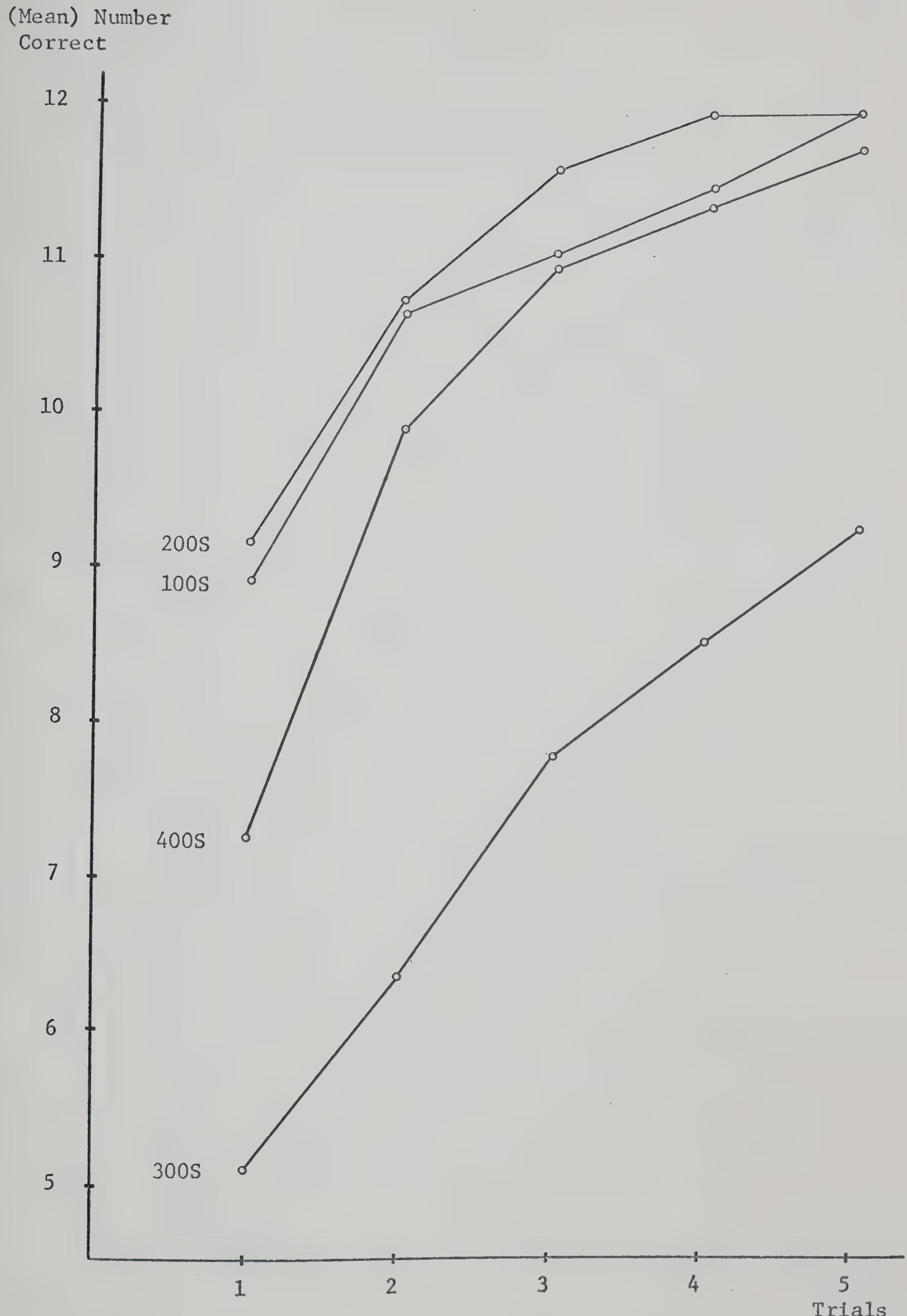


Figure 7. The (Mean) Number Correct on the First Five Trials for the Ordered Lists.

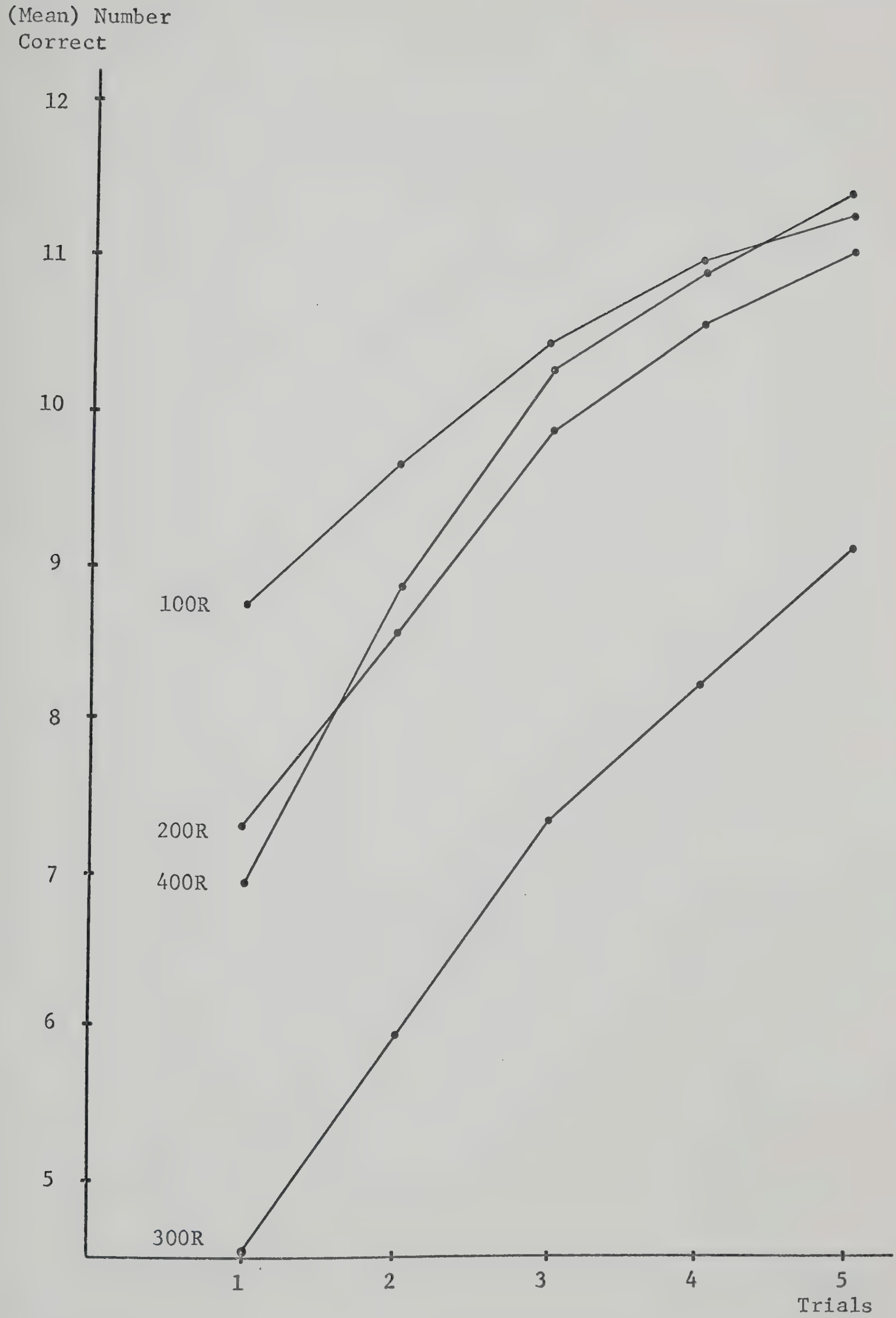


Figure 8. The (Mean) Number Correct on the First Five Trials for the Random (R) Lists.

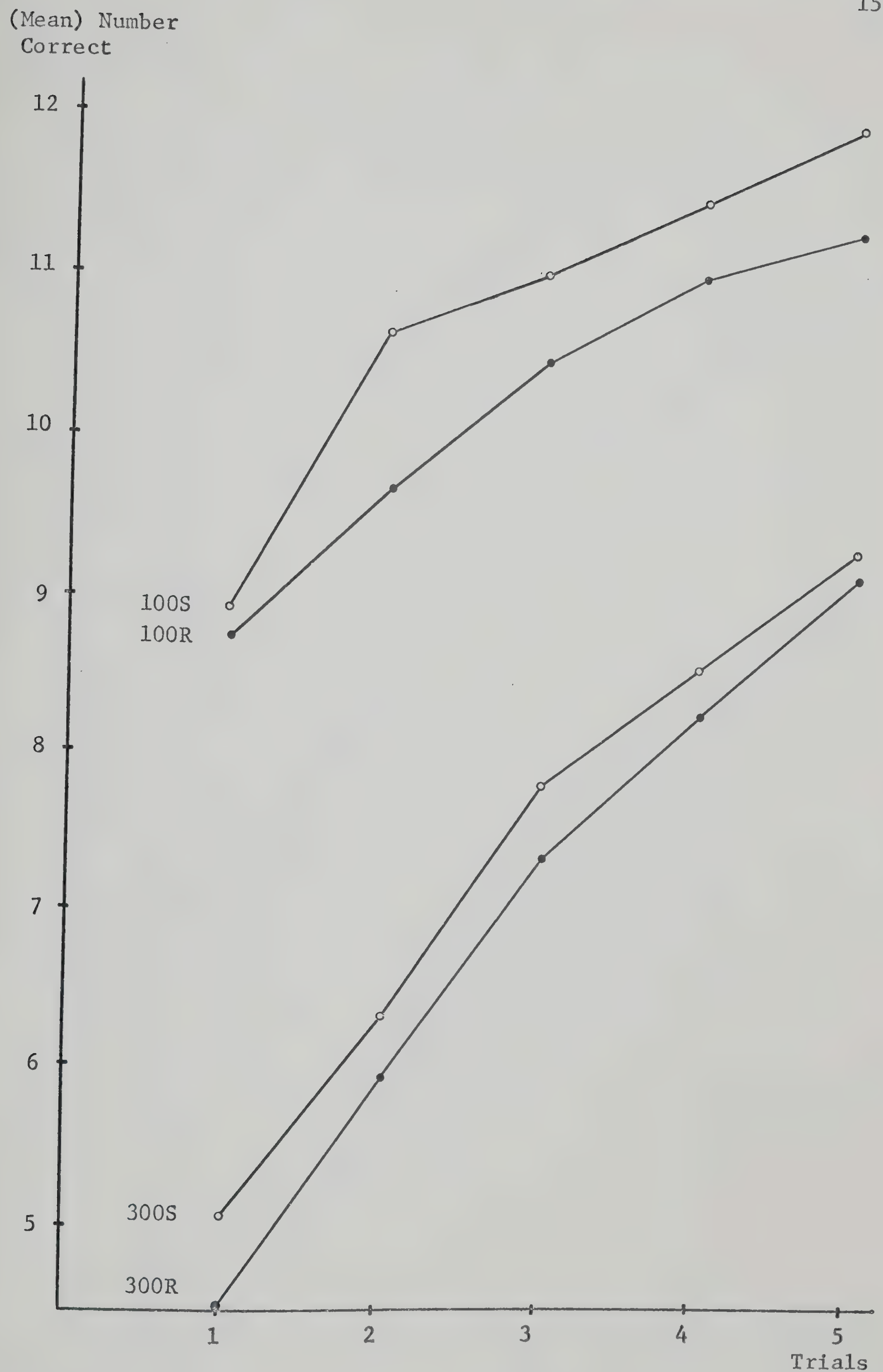


Figure 9. The (Mean) Number Correct on the First Five Trials for Lists 100S, 100R, 300S, and 300R.

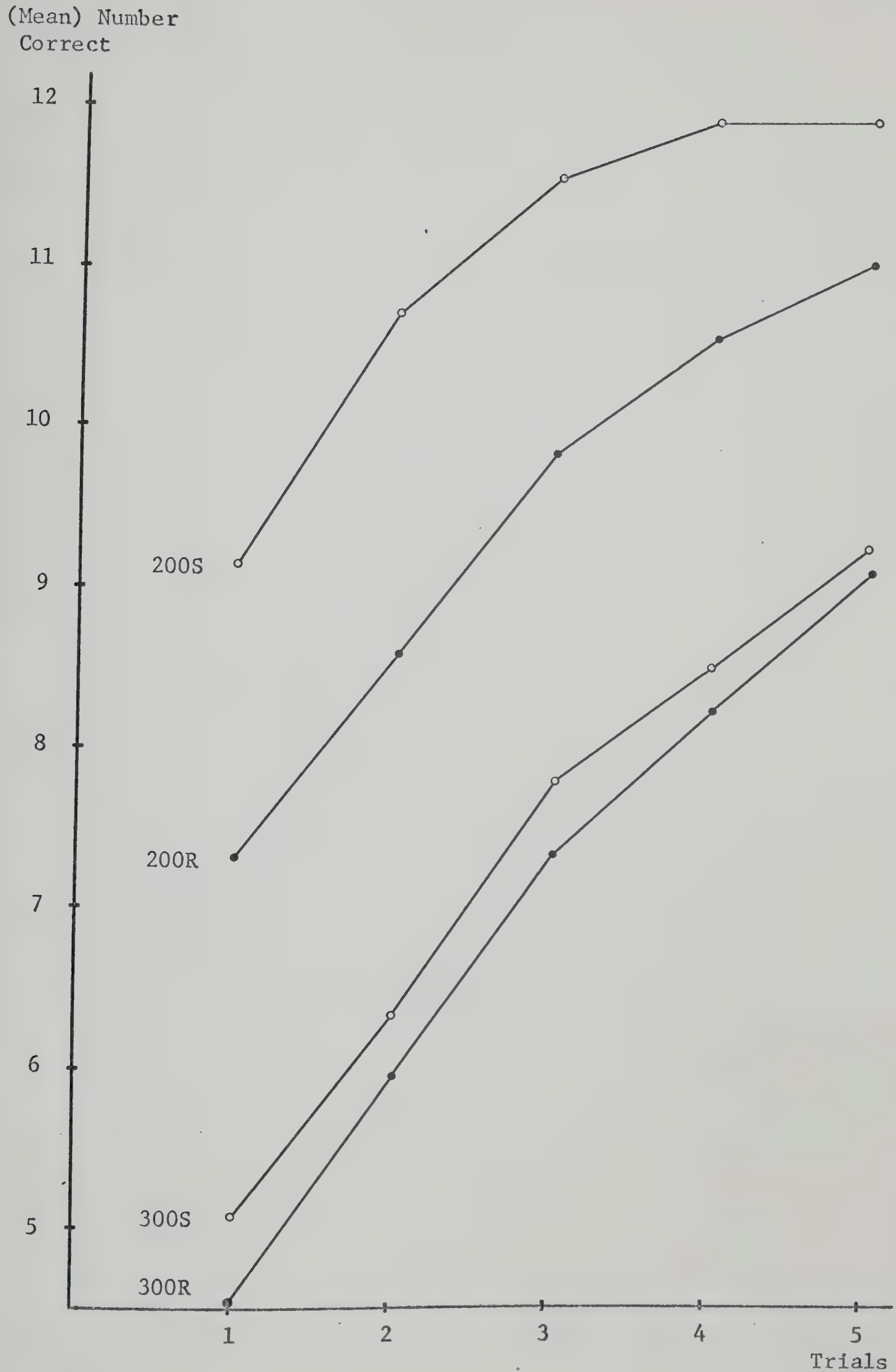


Figure 10. The (Mean) Number Correct on the First Five Trials for Lists 200S, 200R, 300S, and 300R.

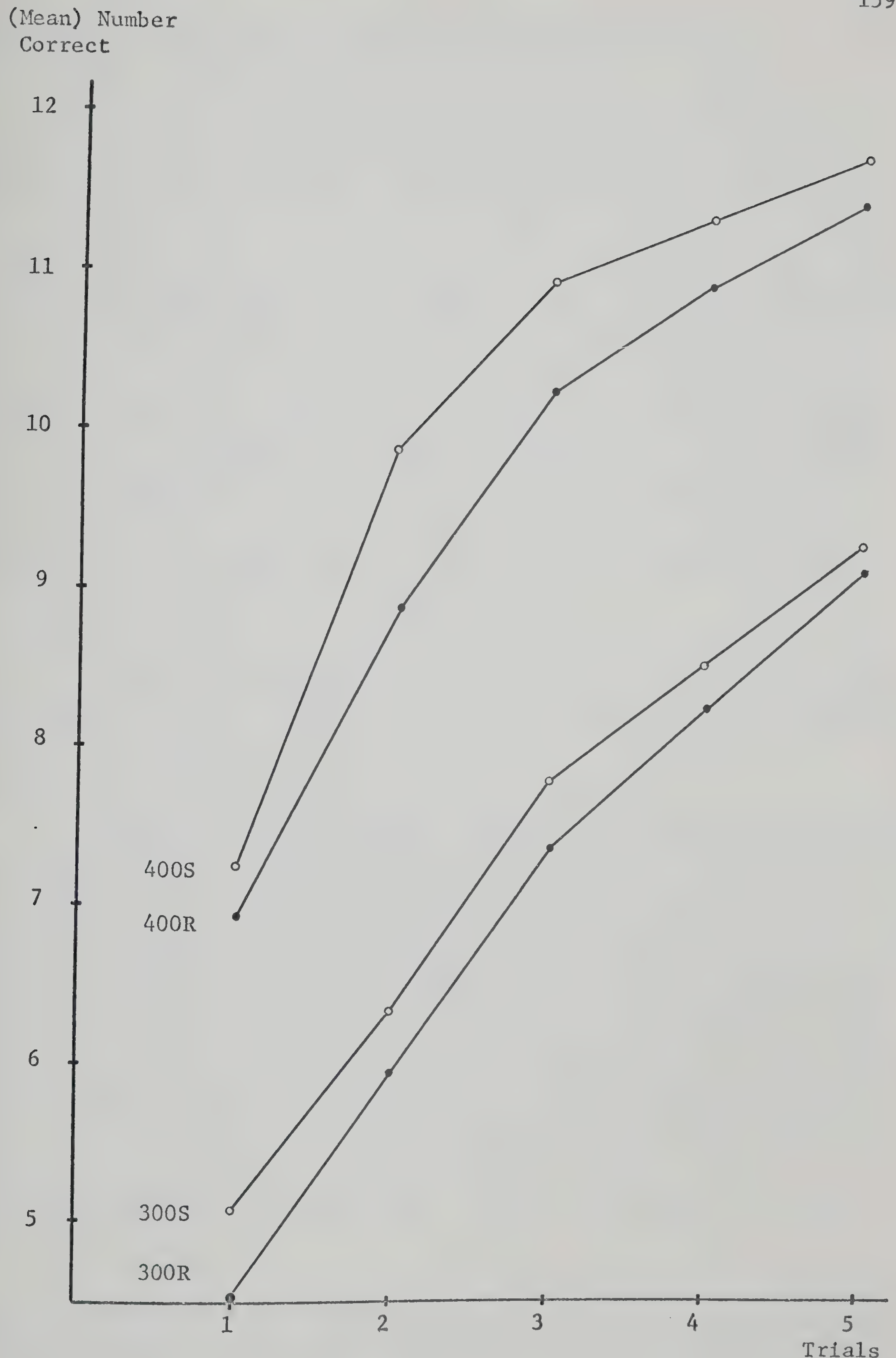


Figure 11. The (Mean) Number Correct on the First Five Trials for Lists 400S, 400R, 300S, and 300R.

inversion, the performance of each sex was examined separately for lists 100 and 200. Figure 12 indicates that there were no significant differences (at 0.05) in sex for the 200 list while order was significant on all trials at the 0.01 level. On the other hand, for the 100 list (Figure 13) order was not significant (at 0.05) on all trials except trial #5 while sex was significant (at 0.05) on all trials except trial #3 and 5. A comparison of the number correct for the 200 list (Table 16) with the performance of the female Ss on the 100S list (Table 19) indicates that the two performances are almost indistinguishable. On the other hand, the performance of the male Ss is consistently inferior on the 100S lists as compared to the 200 lists. Thus the inversion can be explained on the basis of sex differences.[3]

The difference between the 400 list and the others is only significant on trial #1 with list 100 at the 0.01 level, and with list 300 at the 0.0001 level on all five trials.

In summary, the relationship between performance and OP, although not significant at all times, supports the previous trend that as OP increases so does performance; however, the differences in OP have to be relatively large as is the case between list 300 and 400, between list 300 and 100, and between list 300 and 200. The relationship with similarity parallels that of OP except for the 400 list where similarity is lower than for the 300 list but the effect is compensated by the non-similarity based OP in the 400 list. Formal similarity, on the other

---

[3]An analysis of variance was carried out for the three variables of sex by lists by order. Sex proved to be a significant variable for trials to criterion ($p<0.006$), mean number correct per trial ($p<0.02$), and for mean number of errors of omission ($p<0.002$). Mean number of overt errors was not significant, $F<<1$. The summary of analysis of variance is given in Appendix J.

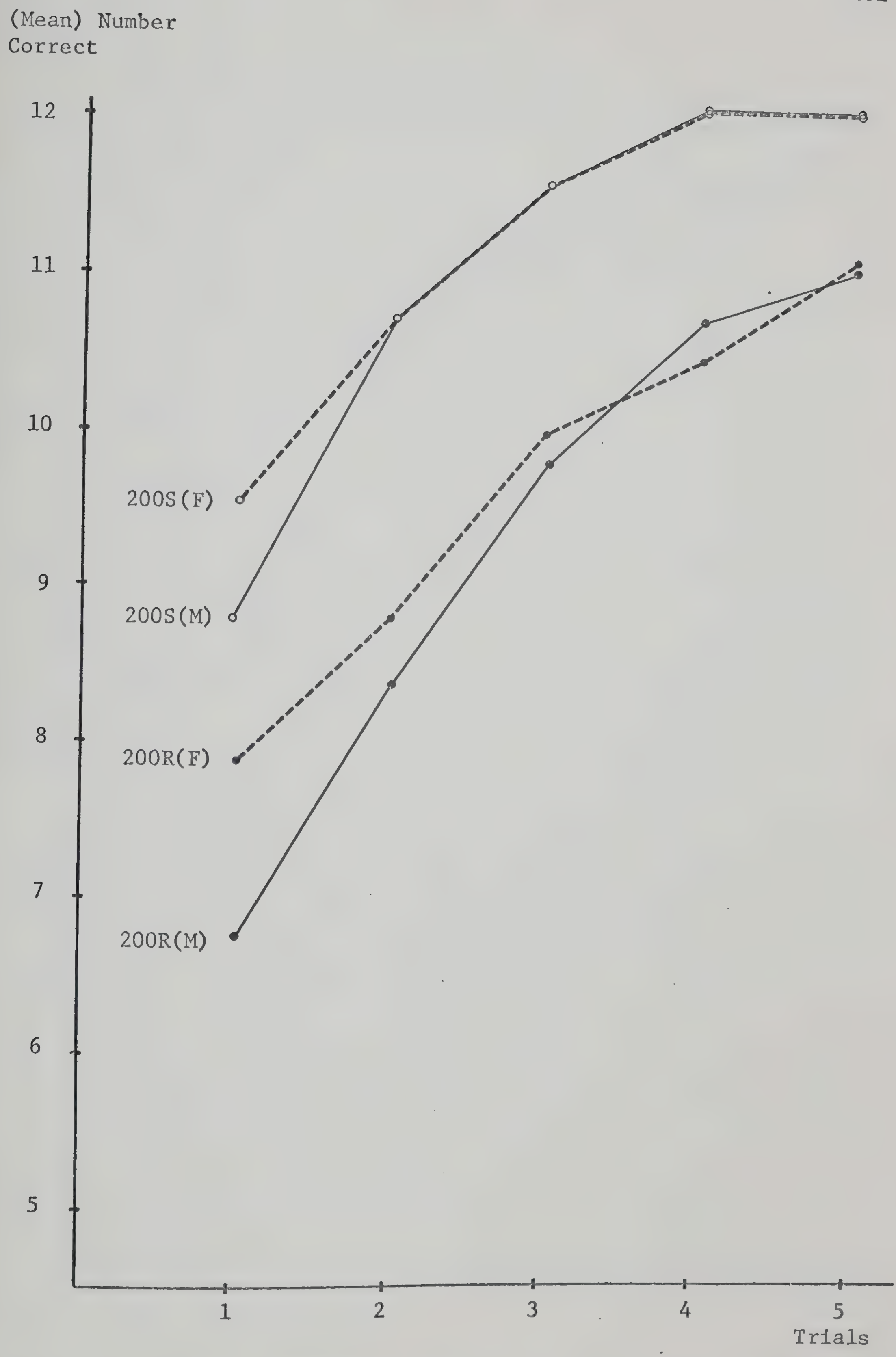


Figure 12. The (Mean) Number Correct on the First Five Trials for Male and Female Subjects for Lists 200S and 200R.

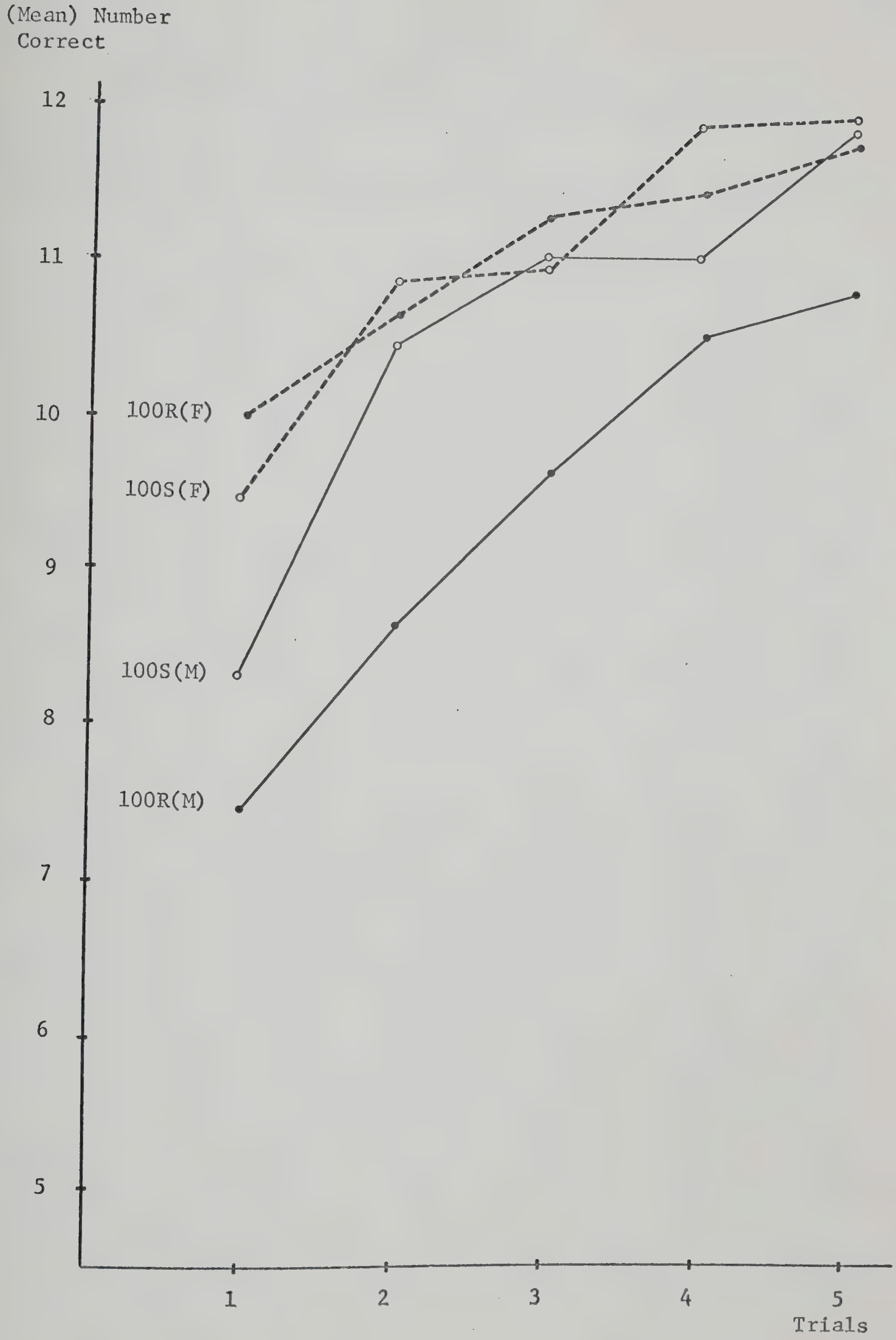


Figure 13. The (Mean) Number Correct on the First Five Trials for Male and Female Subjects for Lists 100S and LOOR.

TABLE 19

THE (MEAN) NUMBER CORRECT IN EACH LIST x ORDER x SEX CONDITION FOR THE FIRST FIVE TRIALS

List, Order & Sex		Trials 1	2	3	4	5
100S	F	9.47	10.87	11.00	11.93	12.00
	M	8.33	10.47	11.07	11.07	11.93
200S	F	9.53	10.73	11.60	11.93	11.93
	M	8.80	10.73	11.60	12.00	12.00
300S	F	5.80	6.73	8.00	8.47	9.73
	M	4.40	5.93	7.60	8.60	8.87
400S	F	7.33	10.27	11.00	11.47	11.87
	M	7.13	9.53	10.93	11.27	11.67
100R	F	10.00	10.67	11.33	11.47	11.80
	M	7.47	8.67	9.67	10.53	10.87
200R	F	7.87	8.80	10.00	10.47	11.13
	M	6.73	8.40	9.80	10.73	11.07
300R	F	4.80	6.00	7.47	8.20	8.87
	M	4.27	5.87	7.27	8.33	9.40
400R	F	7.60	9.27	10.73	11.40	11.80
	M	6.27	8.53	9.87	10.53	11.13

hand, has no consistent effect on performance.

An examination of the number of omission errors as a function of trials showed the same general trend (Figure 14). The overt errors, on the other hand, did not decrease smoothly with increased learning trials but tended to fluctuate from trial to trial (Figure 15). An attempt to clarify the relationship (Figure 16 and 17) showed that while the graphs of the errors for the S lists were reasonably smooth the graphs of the R lists exhibited few errors on the first trial and more errors on the second trial for the 400 and 100 list. But for the 300 list an increase in overt errors occurred on trial 4. It may be conjectured that this difference between the S and R graphs indicates a different learning approach, namely, that the Ss in the R order initially preferred to concentrate on learning a few items making mostly errors of omission while the Ss in the S order attempted to learn the list as a whole. However, the graphical data neither support nor refute such a speculation. The only basis for making it was the observation during the marking of recall protocols that initially the number of overt errors per number attempted was small, but as more items were written down (attempted) the number of errors (overt) increased. This was primarily observed in the 300R list.

In general, however, the examination of the graphs for the overt errors suggests that a much more detailed and specific analysis must be made before the convolutions in the graphs can be adequately explained. The relationship, however, seems very complex and probably depends on a number of different factors such as learning style.

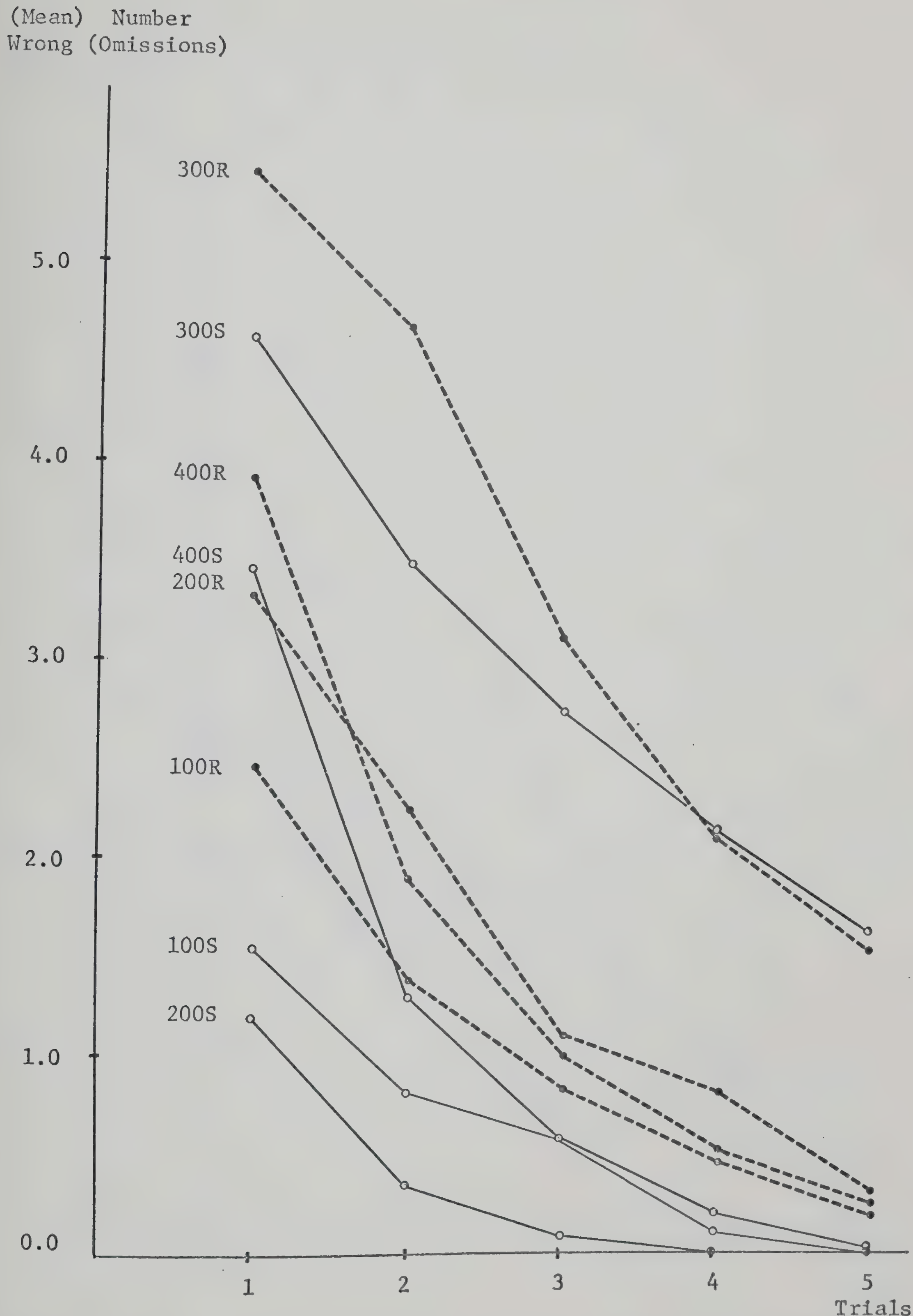


Figure 14. The (Mean) Number Wrong (Errors of Omission) on the First Five Trials for Ordered (S) and Random (R) Lists.

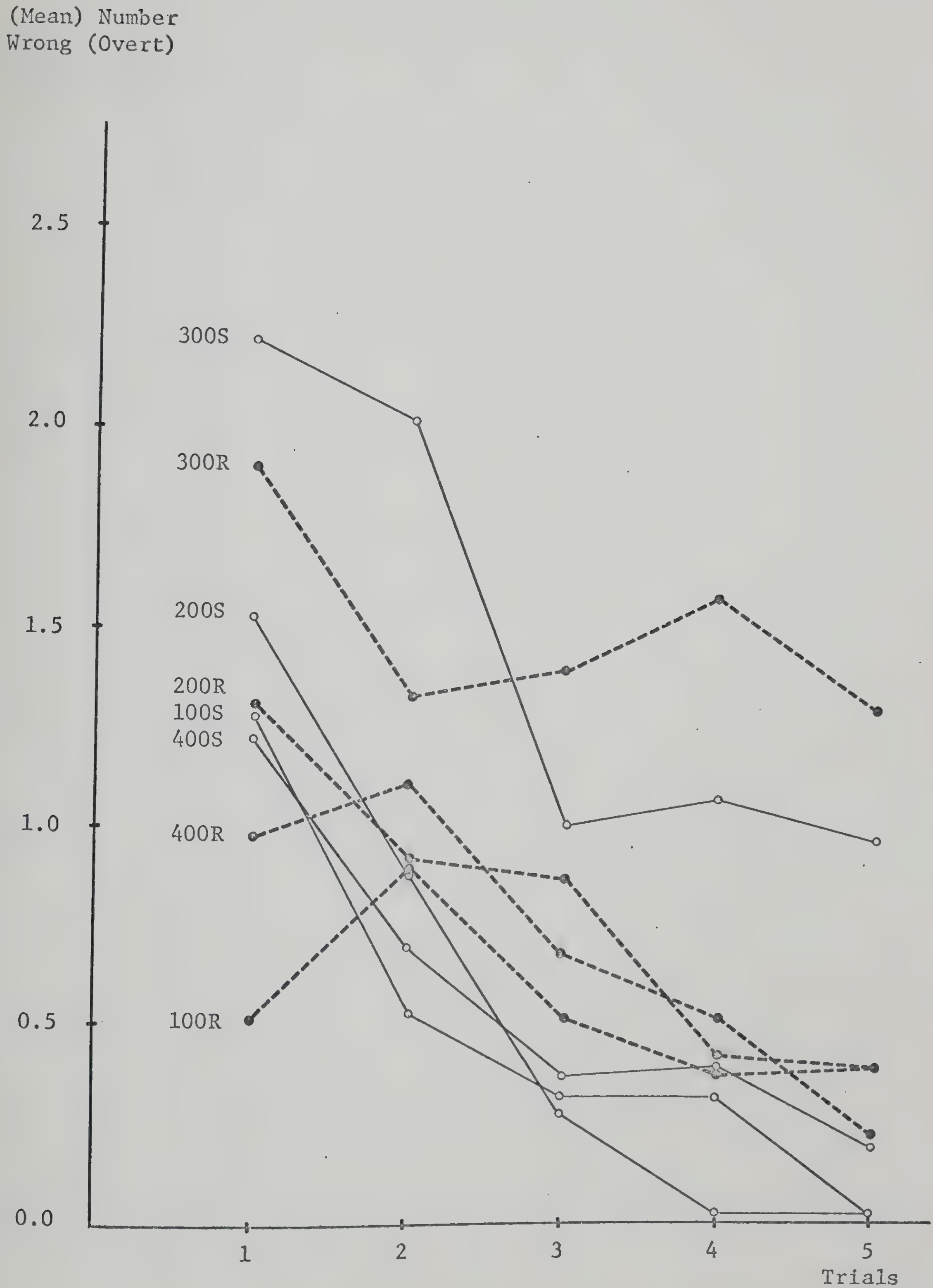


Figure 15. The (Mean) Number Wrong (Overt Errors) on the First Five Trials for the Ordered (S) and Random (R) Lists.

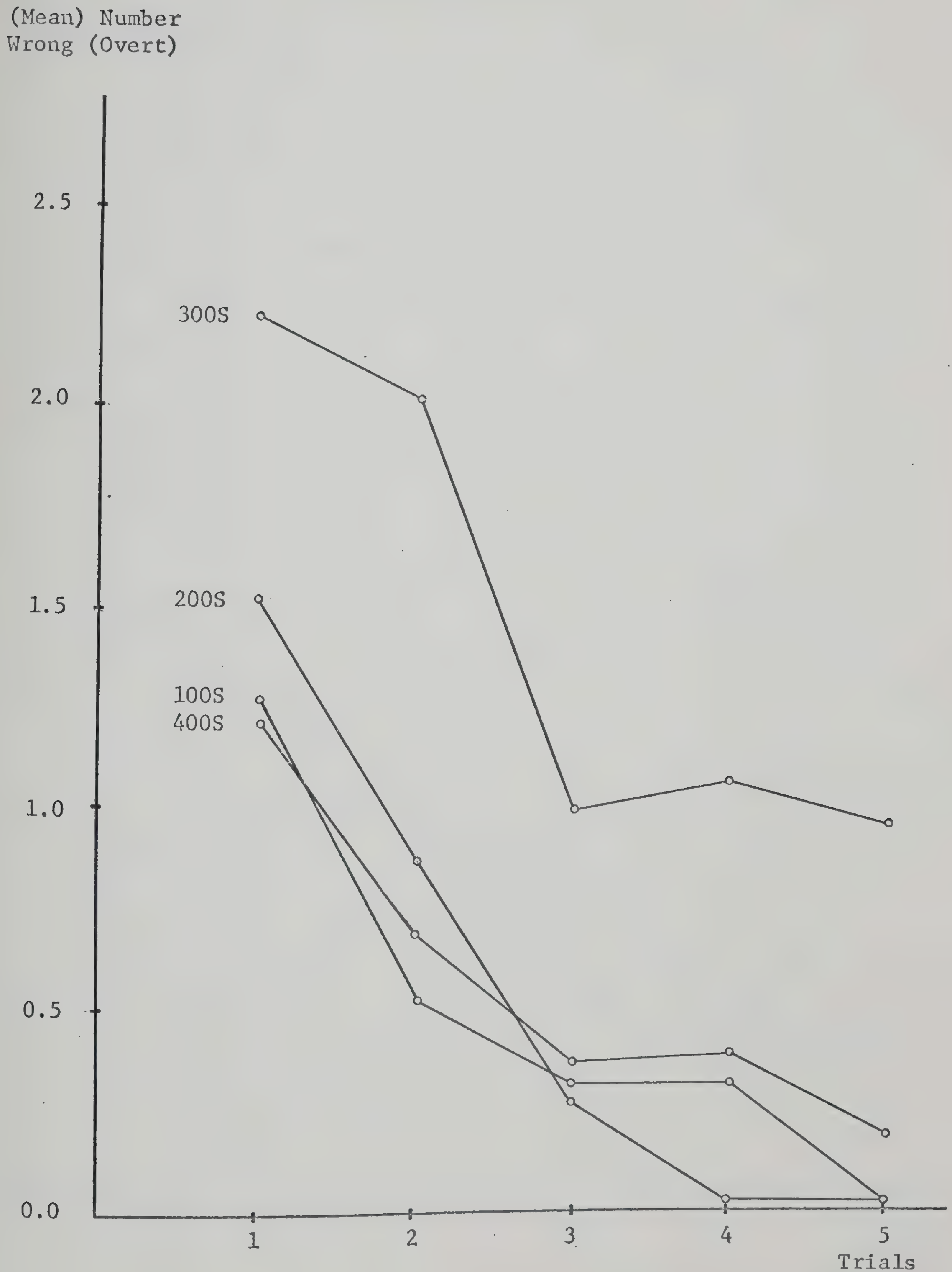


Figure 16. The (Mean) Number Wrong (Overt Errors) on the First Five Trials for Ordered (S) Lists.

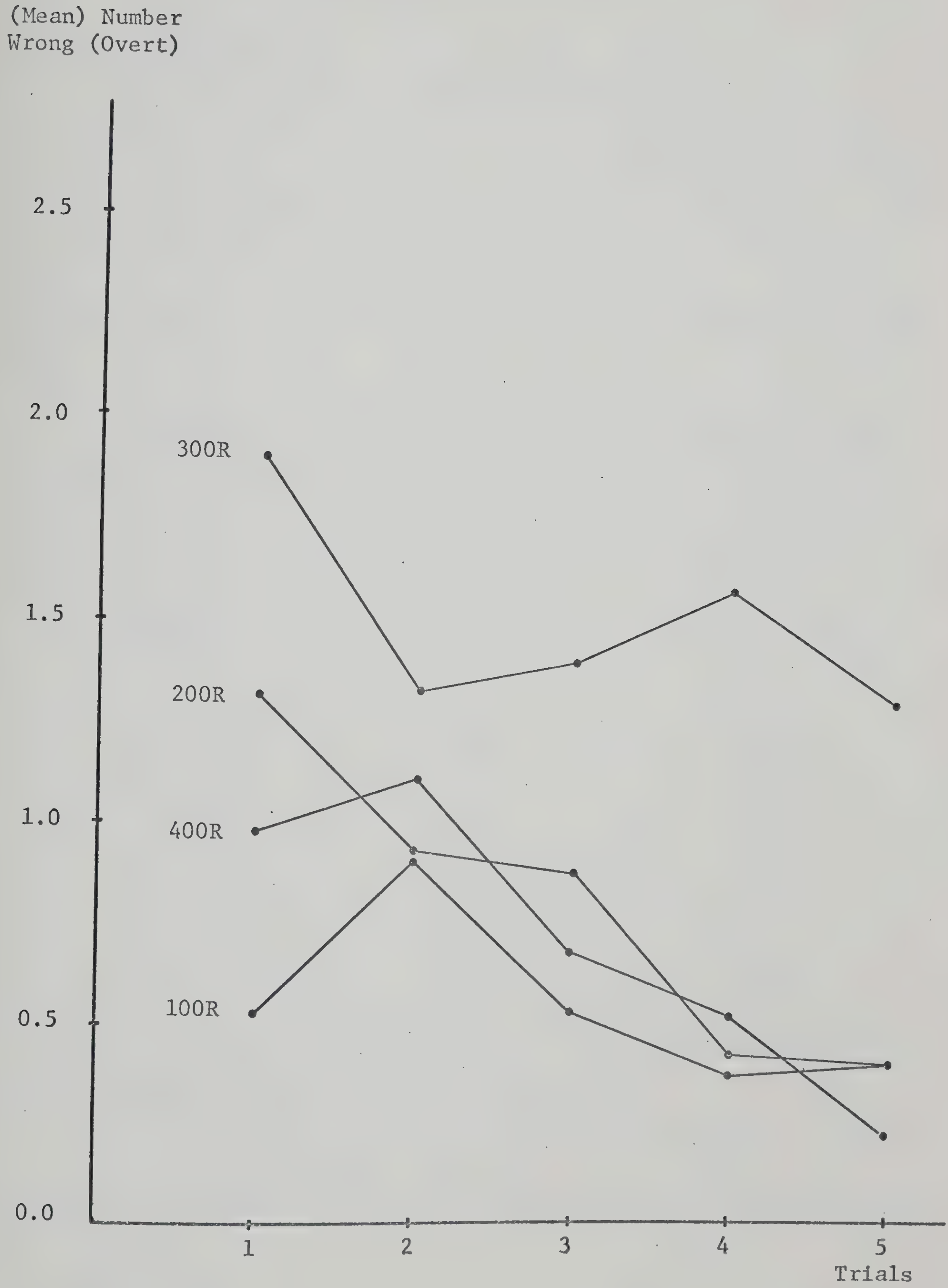


Figure 17. The (Mean) Number Wrong (Overt Errors) on the First Five Trials for Random (R) Lists.

Analysis of Secondary Data

Learning Strategies

One of the purposes of the Questionnaire was to examine the learning strategies used by the Ss. More specifically it was to determine, first of all, the extent to which the Ss used strategies; second, the nature of the strategies; and third, the differences (if any) in the nature and use of strategies for the different list by order treatments.

The table below gives the frequency (and percent) of the Ss reporting the use of one or more learning strategies as well as the number of Ss reporting that they did not use any strategy at all.[4]

Number of Strategies	Number of Ss	Percent of Ss
1	143	59.58
2	63	26.25
3	24	10.00
4	2	0.83
0	8	3.33
Total	240	99.99

It is quite clear that the vast majority of the Ss (96.6%) used some sort of learning strategy to facilitate learning.

The nature and frequency with which these strategies were used is summarized in Table 20. The strategies are divided into two types:

---

[4]The data were compiled on the basis of answers that the Ss gave to questions 1, 3, 4, 6, and 7 on the Questionnaire (Appendix F).

TABLE 20

PERCENT OF SUBJECTS REPORTING THE USE OF A GIVEN STRATEGY

Strategy		100S	100R	200S	200R	300S	300R	400S	400R
1. Similarity Grouping	1.1-1.5[a] (C)	80.0[b]	93.3	96.7	93.3				
	1.6-1.8 (L)					56.7	43.3	60.0 (all 1.6)	70.0 (all 1.6)
	1.9 (S)	20.0	6.7	3.3	6.7	6.7		10.0	6.7
2. Aural Grouping	(L)		3.3			13.3	13.3	6.7	13.3
3. Chaining	(L)		3.3			6.7	10.0		
4. Numerical Grouping	4.1-4.5 (C)					16.7	43.3	33.3	16.7
5. Rhyming and Use of Sound	5.1-5.3 (L)		6.7	3.3	3.3	3.3	3.3	16.7	10.0
6. Alphabetical Sequence	6.1 (C)					16.7		40.0	
	6.2-6.4 (L)	23.3	13.3	3.3	3.3	16.7	3.3	23.3	6.7
7. Miscellaneous	7.2-7.7 (C)	3.3	3.3	6.7		13.3	16.7	3.3	13.3
	7.1,7.8,7.9 (L)	10.0	13.3		3.3	13.3	16.7	10.0	6.7

[a]Strategy numbers refer to Table #7.

[b]Percentages total more than 100% per column because Ss could use more than one strategy.

comprehensive and limited. As the name implies, 'limited strategies' are somewhat restricted in usefulness; nevertheless, their use by the Ss supports the hypothesis that Ss do attempt to structure the StOs in a list.[5]

From Table 20 it can be seen that there are basic differences between the strategies used by the Ss learning the 100-200 lists and the strategies used by the Ss learning the 300-400 lists. The major differences, of course, confirm the basic structure of the lists in terms of similarity. That is, the predominant use of similarity grouping as the major and almost exclusive strategy in the 100-200 lists suggests that they are composed of highly similar items that could be structured into a pattern. On the other hand, the use of numerical grouping as a major comprehensive strategy in the 300-400 lists indicates that they are relatively unstructurable in terms of similarity. The numerical grouping strategy is characterized by a lack of a grouping basis. Thus it may be inferred that it would be used only in lists that did not offer a better strategy.

The effect of item order on strategy may be observed in the case of alphabetical order. In the S condition of the 300 and 400 lists 16.7% and 40.0% of the Ss respectively reported the use of alphabetical order; however, under the R conditions not one subject indicated the use of this strategy. This suggests that a given item may mask a strategy and prevent it from being used.

Coding by Spelling or Pronunciation

Question # 8 of the Questionnaire required that the Ss list

---

[5]For a discussion of 'limited strategies' see Chapter 6, p.133.

all the CVC's and indicate if they learned a given CVC by spelling it or by pronouncing it.[6] The results indicated that all lists were learned by both methods. The data are presented in Table 21 in percent form. (The raw data are given in Table 24, Appendix I).

For the 100 list four trigrams (XUV, ZOH, ZOS, and ZOJ) were spelled by at least two thirds of the Ss. The exact percentages ranged from 66.7% for XUV to 73.3% for ZOJ. Eight trigrams were spelled by approximately half of the Ss (percentages ranged from 41.7% for HOZ and JOZ to 55.0% for KUV). In all cases the number of Ss spelling a given trigram exceeded the number that pronounced it; however, at times the difference was less than 5%.

In the 200 list, the same pattern was repeated. Four trigrams (XUV, ZOH, QIF, QIJ) were spelled by 60.0% or more of the Ss. The exact percentages ranged from 60.0% for XUV and ZOH to 70.0% for QIJ. Eight trigrams were spelled by approximately half of the Ss (46.7% for VUK and SOZ to 58.3% for ZOS). In all cases the CVC's were spelled by more Ss.

In the 300 list five CVC's were pronounced by at least 60% of the Ss. The trigrams were the following: ZIV, 61.7%; VOF, 63.3%; VUK and HUZ, 71.7% each; and FUX, 78.3%. Five other trigrams (KUJ, XIF, QIJ, QOS, and XOQ) were spelled by at least 60% of the Ss (61.7% to 75.0%). The two remaining trigrams were split half and half. JIH was spelled by 51.7% of the Ss while SOZ was pronounced by 58.3% of the Ss.

---

[6] For some trigrams the Ss either did not specify the code or indicated that they used both codes. The percent frequency of such instances for each trigram (designated as U in Table 21) does not exceed 20%, for any given trigram except for four CVC's in the 200 list (FIQ, XUV, JIQ, and SOZ). The latter, however, did not exceed 30%. The actual distribution (number of Ss) is given in Table 24, Appendix I.

TABLE 21

THE PERCENTAGE OF SUBJECTS REPORTING THE CODING OF NONSENSE SYLLABLES BY PRONUNCIATION (P) OR BY SPELLING (S)[a]

List	Code[b]	VUK	VUT	VUX	KUV	TUV	XUV	ZOH	ZOJ	ZOS	HOZ	JOZ	SOZ
100	P	40.0	43.3	35.0	28.3	33.3	15.0	18.3	11.7	18.3	41.7	43.3	33.3
	S	48.0	48.3	53.3	55.0	51.7	66.7	66.7	73.3	68.3	41.7	46.7	53.3
	U	11.7	8.3	11.7	16.7	15.0	18.3	15.0	15.0	13.3	16.7	10.0	13.3

List	Code[b]	VUK	QIF	VUX	KUV[c]	FIQ	XUV	ZOH	QIJ	ZOS	HOZ[c]	JIQ	SOZ
200	P	33.3	31.7	36.7	31.7	18.3	11.7	20.0	23.3	31.7	40.0	16.7	25.0
	S	46.7	61.7	55.0	51.7	55.0	60.0	60.0	70.0	58.3	48.3	56.7	46.7
	U	20.0	6.7	8.3	16.7	26.7	28.3	20.0	6.7	10.0	11.7	26.7	28.3

[a]The raw data (the number of Ss) are given in Table 24, Appendix I.

[b]Key: P - pronounced
S - spelled
U - unknown (Ss did not specify code)

[c]For breakdown of percentages into Sex x Order see Table 22, p. 176.

TABLE 21 (continued)[a]

List	Code[b]	VUK	FUX	HUZ	JIH	KUJ	QOS	VOF	QIJ	XIF	XOQ	ZIV	SOZ
300	P	71.7	78.3	71.7	40.0	26.7	21.7	63.3	21.7	25.0	16.7	61.7	58.3
	S	20.0	13.3	20.0	51.7	61.7	66.7	28.3	65.0	63.3	75.0	28.3	31.7
	U	8.3	8.3	8.3	8.3	11.7	11.7	8.3	13.3	11.7	8.3	10.0	10.0

List	Code[b]	VUK	BOF	CEH[c]	GYX	LAJ[c]	MYP	NID	QIJ	ZOS	QUR	TEH[c]	VAW
400	P	70.0	78.3	50.0	23.3	50.0	41.7	81.7	33.3	68.3	35.0	53.3	66.7
	S	28.3	16.7	46.7	73.3	43.3	55.0	15.0	61.7	26.7	58.3	43.3	30.0
	U	1.7	5.0	3.3	3.3	6.7	3.3	3.3	5.0	5.0	6.7	3.3	3.3

[a]The raw data (the number of Ss) are given in Table 24, Appendix I.

[b]Key: P - pronounced
S - spelled
U - unknown (Ss did not specify code)

[c]For breakdown of percentages into Sex x Order see Table 22, p. 176.

In the 400 list the distribution was more complex. The largest group consisted of five trigrams that were pronounced by at least two thirds of the <u>Ss</u>. These were VAW, 66.7%; ZOS, 68.3%; VUK, 70.0%; BOF, 78.3%; and NID, 81.7%. Two trigrams (QIJ and GYX) were spelled by 61.7% and 73.3% of the <u>Ss</u> (respectively). Two trigrams (CEH and TEH) were treated differently by the female <u>Ss</u> in the R list order condition (Table 22). 66.7% spelled CEH and 60.0% spelled TEH. In all the other treatments (S and R, male; S, female) approximately 60.0% of the <u>Ss</u> pronounced each of these trigrams. Another CVC that was treated differentially was LAJ. In the S condition 73.3% of the female <u>Ss</u> pronounced it while 66.7% of the male <u>Ss</u> spelled it. In the R condition 53.3% of the male <u>Ss</u> pronounced it while 53.3% of the female <u>Ss</u> spelled it.

In summary, the results indicated that whereas the 100 and 200 lists were primarily spelled the 300 and 400 lists were primarily half pronounced and half spelled. Furthermore, it was found that some CVC's were treated differently under different list order conditions, and that there were sex differences for three trigrams.

<u>Context Effect</u>. A number of trigrams appeared in more than one list; thus it was possible to observe any effects due to list context on the coding mode used by the <u>Ss</u> (Table 21, p.173 and Table 22, p.176).

Eight trigrams were used in both the 100 and 200 lists. An examination of the frequency with which they were spelled in each list (Table 21) indicated that they were on the whole coded the same way. The only difference occurred for KUV and HOZ. In the 100 list they were spelled by 55.0% and 41.7% of the <u>Ss</u> respectively; however, in the 200 list under R item order they tended to be pronounced. The largest

TABLE 22

THE PERCENTAGE OF MALE AND FEMALE SUBJECTS REPORTING THE CODING OF NONSENSE SYLLABLES BY PRONUNCIATION (P) OR BY SPELLING (S) FOR EACH LIST x ORDER TREATMENT

List	Order	Sex	Code[b]	VUK	VUT	VUX	KUV	TUV	XUV	ZOH	ZOJ	ZOS	HOZ	JOZ	SOZ
100	S	F	P	33.3	33.3	33.3	40.0	40.0	13.3	20.0	20.0	26.7	40.0	40.0	40.0
			S	60.0	60.0	60.0	53.3	46.7	73.3	80.0	80.0	73.3	46.7	46.7	46.7
			U	6.7	6.7	6.7	6.7	13.3	13.3				13.3	13.3	13.3
		M	P	53.3	53.3	46.7	20.0	26.7	6.7	13.3	6.7	20.0	40.0	46.7	46.7
			S	40.0	40.0	46.7	46.7	46.7	60.0	60.0	60.0	53.3	40.0	46.7	46.7
			U	6.7	6.7	6.7	33.3	26.7	33.3	26.7	33.3	26.7	20.0	6.7	6.7
	R	F	P	33.3	40.0	20.0	20.0	26.7	13.3	20.0	6.7	6.7	46.7	46.7	20.0
			S	53.3	53.3	60.0	66.7	66.7	73.3	60.0	80.0	80.0	40.0	46.7	66.7
			U	13.3	6.7	20.0	13.3	6.7	13.3	20.0	13.3	13.3	13.3	6.7	13.3
		M	P	40.0	46.7	40.0	33.3	40.0	26.7	20.0	13.3	20.0	40.0	40.0	26.7
			S	40.0	40.0	46.7	53.3	46.7	60.0	66.7	73.3	66.7	40.0	46.7	53.3
			U	20.0	13.3	13.3	13.3	13.3	13.3	13.3	13.3	13.3	20.0	13.3	20.0

[a]The raw data (the number of Ss) are given in Table 25, Appendix I.

[b]Key: P - pronounced
S - spelled
U - unknown (Ss did not specify code)

TABLE 22 (continued)[a]

List	Order	Sex	Code[b]	VUK	QIF	VUX	KUV	FIQ	XUV	ZOH	QIJ	ZOS	HOZ	JIQ	SOZ
200	S	F	P	33.3	26.7	33.3	6.7			13.3	13.3	26.7	20.0		13.3
			S	66.7	73.3	66.7	86.7	93.3	93.3	86.7	86.7	73.3	73.3	86.7	73.3
			U				6.7	6.7	6.7				6.7	13.3	13.3
		M	P	40.0	33.3	33.3	13.3	13.3	13.3	33.3	33.3	33.3	20.0	20.0	20.0
			S	46.7	53.3	53.3	53.3	53.3	53.3	60.0	60.0	60.0	46.7	53.3	53.3
			U	13.3	13.3	13.3	33.3	33.3	33.3	6.7	6.7	6.7	33.3	26.7	26.7
	R	F	P	26.7	40.0	33.3	60.0	26.7	13.3	13.3	26.7	40.0	66.7	26.7	40.0
			S	40.0	60.0	53.3	33.3	40.0	60.0	53.3	73.3	46.7	33.3	40.0	26.7
			U	33.3		13.3	6.7	33.3	26.7	33.3		13.3		33.3	33.3
		M	P	33.3	26.7	46.7	46.7	33.3	20.0	20.0	20.0	26.7	53.3	20.0	26.7
			S	33.3	60.0	46.7	33.3	33.3	33.3	40.0	60.0	53.3	40.0	46.7	33.3
			U	33.3	13.3	6.7	20.0	33.3	46.7	40.0	20.0	20.0	6.7	33.3	40.0

[a]The raw data (the number of Ss) are given in Table 25, Appendix I.

[b]Key: P - pronounced
S - spelled
U - unknown (Ss did not specify code)

TABLE 22 (continued)[a]

List	Order	Sex	Code[b]	VUK	FUX	HUZ	JIH	KUJ	QOS	VOF	QIJ	XIF	XOQ	ZIV	SOZ
300	S	F	P	73.3	93.3	86.7	46.7	13.3	26.7	73.3	6.7	20.0	6.7	66.7	66.7
			S	26.7	6.7	13.3	53.3	80.0	66.7	26.7	93.3	80.0	93.3	33.3	33.3
			U					6.7	6.7						
		M	P	80.0	86.7	53.3	40.0	33.3	20.0	60.0	20.0	20.0	20.0	60.0	60.0
			S	13.3	6.7	26.7	53.3	53.3	60.0	33.3	53.3	73.3	66.7	26.7	20.0
			U	6.7	6.7	20.0	6.7	13.3	20.0	6.7	26.7	6.7	13.3	13.3	20.0
	R	F	P	66.7	73.3	73.3	53.3	26.7	20.0	60.0	33.3	33.3	20.0	60.0	53.3
			S	33.3	26.7	26.7	46.7	66.7	80.0	33.3	60.0	60.0	80.0	40.0	46.7
			U					6.7		6.7	6.7	6.7			
		M	P	66.7	60.0	73.3	20.0	33.3	20.0	60.0	26.7	26.7	20.0	60.0	53.3
			S	6.7	13.3	13.3	53.3	46.7	60.0	20.0	53.3	40.0	60.0	13.3	26.7
			U	26.7	26.7	13.3	26.7	20.0	20.0	20.0	20.0	33.0	20.0	26.7	20.0

[a]The raw data (the number of Ss) are given in Table 25, Appendix I.

[b]Key: P - pronounced
S - spelled
U - unknown (Ss did not specify code)

TABLE 22 (continued)[a]

List	Order	Sex	Code[b]	VUK	BOF	CEH	GYX	LAJ	MYP	NID	QIJ	ZOS	QUR	TEH	VAW
400	S	F	P	73.3	80.0	60.0	33.3	73.3	46.7	80.0	46.7	73.3	53.3	66.7	73.3
			S	20.0	13.3	33.3	66.7	20.0	53.3	13.3	53.3	20.0	40.0	26.7	20.0
			U	6.7	6.7	6.7		6.7		6.7		6.7	6.7	6.7	6.7
		M	P	60.0	73.3	53.3	20.0	33.3	33.3	80.0	40.0	73.3	40.0	53.3	60.0
			S	40.0	26.7	46.7	80.0	66.7	66.7	20.0	60.0	26.7	60.0	46.7	40.0
			U												
	R	F	P	60.0	73.3	26.7		40.0	33.3	86.7	6.7	53.3	20.0	33.3	60.0
			S	40.0	20.0	66.7	100.0	53.3	66.7	13.3	86.7	40.0	73.3	60.0	33.0
			U		6.7	6.7		6.7			6.7	6.7	6.7	6.7	6.7
		M	P	86.7	86.7	60.0	40.0	53.3	53.3	80.0	40.0	73.3	26.7	60.0	73.3
			S	13.3	6.7	40.0	46.7	33.3	33.3	13.3	46.7	20.0	60.0	40.0	26.7
			U		6.7		13.3	13.3	13.3	6.7	13.3	6.7	13.3		

[a]The raw data (the number of Ss) are given in Table 25, Appendix I.

[b]Key: P - pronounced
S - spelled
U - unknown (Ss did not specify code)

difference occurred for the female Ss (Table 22). Thus, while in the S condition 86.7% of the females spelled KUV, in the R condition 60.0% of them pronounced it. A similar inversion occurred for HOZ.

Two trigrams, ZOS and SOZ, also were used in the 400 and 300 lists respectively. An examination of the frequency with which they were spelled (Table 21) indicated that whereas in the 100 and 200 lists they were spelled on the average by more than half of the Ss, in the 400 and 300 lists they were primarily pronounced (68.3% of the Ss pronounced ZOS in the 400 list, and 58.3% of the Ss pronounced SOZ in the 300 list).

The greatest effect, however, was observed for VUK which was used in all four lists (Table 21). In the 100 and 200 lists VUK was pronounced by less than half of the Ss (40.0% and 33.3% respectively); however, in the 300 and 400 lists it was pronounced by almost three quarters of the Ss (71.7% and 70.0% respectively).

These findings seem to indicate that the context of a trigram may play an important role as to the way in which it will be coded. Consequently, a trigram may not be describable as either pronounceable or spellable without some reference being made as to the other CVC's that are to appear with it in a given list.

Convenience of Item Order in the Lists

One of the minor purposes of the Questionnaire was to determine if the Ss found the given item order inconvenient. The responses to this question (#9) are shown in Table 23.

The data indicate that more students found the R condition inconvenient than the S condition. In general between 6.7% and 16.7% of the Ss found the S condition inconvenient, while between 30.0% and 50.0% of the Ss found the R condition inconvenient. In terms of lists,

TABLE 23

FREQUENCY DISTRIBUTION OF SUBJECTS
FINDING ITEM ORDER INCONVENIENT

List & Order	Female Ss	Male Ss	Both (F & M)
100 S	6.7%	13.3%	10.0%
100 R	60.0%	40.0%	50.0%
200 S	0.0%	33.3%	16.7%
200 R	53.3%	33.3%	43.3%
300 S	0.0%	33.3%	16.7%
300 R	40.0%	26.7%	33.3%
400 S	13.3%	0.0%	6.7%
400 R	46.7%	13.3%	30.0%

SUB-TABLES

List & Order	F M	List & Order	F M
100 S	6.7% < 13.3%	100 R	60.0% > 40.0%
200 S	0.0% < 33.3%	200 R	53.3% > 33.3%
300 S	0.0% < 33.3%	300 R	40.0% > 26.7%
400 S	13.3% > 0.0%	400 R	46.7% > 13.3%

the R condition was most inconvenient for the 100 and 200 lists and less so for the 300 and 400 lists. That is, the lists with higher degrees of similarity (and OP) tended to be more inconvenient under R condition than the low similarity (and OP) lists.

The same relationship held for each sex; however, the female Ss in all lists rated the R condition more frequently as inconvenient than the male Ss. That is, while 40.0%, 33.3%, 26.7% and 13.3% of the male Ss indicated that the R condition of the 100, 200, 300, and 400 lists respectively was inconvenient, 60.0%, 53.3%, 40.0%, and 46.7% of the female Ss indicated inconvenience for the R order in these lists respectively. The second Sub-Table on p. 181 presents this data in summarized form.

For the S item order, on the other hand, the male Ss reported inconvenience more often than the female Ss except in the case of the 400S list where again more female Ss (13.3%) registered inconvenience than male Ss (0.0%).

Attempts to relate the frequency of reported inconvenience for the sexes to performance were not successful. However, overall reaction was found to be related to general performance. That is, R order was inconvenient and it resulted in the poorer performance relative to the more preferred S order. Within each order, however, there was no consistent relationship since, for example, the 100R list netted 50.0% inconvenient reaction as compared to 33.3% for the 300R lists; yet the 100R list resulted in better performance than the 300R list.

Summary

In general, the findings may be taken as supporting all four

hypotheses. More specifically they may be summarized in the following points.

1. It was found that a high degree of overall list similarity (OP defined) tended to facilitate learning. In terms of the number of trials to criterion, the lists high and moderate in similarity were learned significantly faster than lists with a limited degree of similarity ($p<0.00001$). Furthermore, the high and moderate similarity lists were learned better than lists consisting of almost all different items; however, the difference did not reach significance at the 0.05 level. The same results were obtained for other performance measures: mean number correct per trial, mean number of overt errors per trial, and mean number of omission errors per trial. In the last measure the difference between the very high similarity list and the different item list reached significance at the 0.05 level ($p<0.03$) with the very high similarity list yielding the superior performance. This may be taken as support for Hypothesis IV.

2. It was found that organization of the items in the lists tended to facilitate recall. The performance of the Ss in the S conditions was significantly superior to that of the Ss in the R conditions in terms of the number of trials to criterion ($p<0.0002$), and mean number correct per trial ($p<0.03$). This may be taken as support for Hypothesis III.

3. The effect of formal similarity on learning was observed to be not systematic. The 100, 200, and 300 lists were equally high in formal similarity but resulted in superior and inferior performance respectively. (The differences between 100 and 200 were not significant).

4. The nature of the strategies used reflected the basic structure of the lists. The high-moderate similarity lists were organized on the basis of similarity; the low similarity lists were organized primarily in terms of numerical grouping and alphabetical sequence. Similarity grouping in the low similarity lists occurred only as a limiting strategy. However, it was used extensively (43.3% to 70.0%). This may be taken as support for Hypothesis II that similarity has a potential to serve as a basis for organization.

5. The almost exclusive use of similarity as an organizational tool in the 100-200 lists and its use as the most frequent limiting strategy in the 300-400 lists indicates that similarity may serve as a basis for organization. This may be taken to mean that similarity has OP which would support Hypothesis II.

6. The large percentage of Ss (96.7%) reporting the use of one or more learning strategies indicates that Ss do not primarily learn by rote. This may be taken as support for Hypothesis I.

7. The effect of item order on the use of strategies was observed in the case of the alphabetical-sequence strategy. In the R condition this strategy was not used while in the S condition it was used to some extent (16.7% to 40.0%). This may be taken as indicating that a given item order may mask the perception of a given strategy and therefore preclude its use. Such an effect was not observed with the more prominent similarity grouping strategy in the high-medium similarity lists.

8. The convenience of a given item order for the Ss was found to be related at a gross level to performance. The R item order was consis-

tently reported as inconvenient by more Ss than the S item order. The maximum difference occurred in the case of the 100 list (40.0%). In each list there was an indication of a sex difference. The female Ss generally reported the R order more frequently as inconvenient than the male Ss. In the S condition the reverse was true. These sex differences, however, could not be related in a systematic way to performance.

9. One set of lists (100-200) was primarily coded by spelling the items; another set (300-400) was coded approximately half by spelling and half by pronunciation. However, the performance on the first set of lists was superior to that on the second set. Therefore, pronounceability *per se* cannot be considered as leading to a better performance. The existence of a context effect indicated that the coding of a CVC as a spellable or pronounceable unit depends on the other items in the list.

In conclusion, the data indicate that Ss are active organizers who attempt to structure the material that is to be learned; that organization facilitates learning; that similarity may be successfully used as a basis of organization; and that a high degree of similarity (OP defined) facilitates learning under free learning conditions.

Chapter 8

DISCUSSION

The present study defined overall list similarity in terms of organizational potential (OP) and therefore the findings will be discussed in relation to the three parameters of OP.

The difference in performance on the 100, 200, and 300 lists may be characterized in terms of the difference in the parameters of similarity based organizational potential (OP). The 100 list provides very few alternatives for grouping; that is, there is only one major structuring scheme available: VU, ZO, mirror-images of VU, and mirror-images of ZO. The 200 list provides two competing alternatives. The items could be grouped either into mirror-image pairs (VUK-KUV), or on the basis of common letters (VUK, VUX). An examination of the recall protocols of the Ss showed that they tended to vacillate between these two groupings from trial to trial. Nevertheless, this did not seem to affect the overall performance of the 200 group. The reason for this may be that both types of groupings could be combined into one more comprehensive grouping, that is, VUK-KUV
VUX-XUV.

In fact, on the Questionnaire 28.3% of the Ss reported using such a scheme. In the 100 list no such vacillation was observed. From the very beginning Ss selected one mode and used it consistently. Although technically speaking, there were two modes available, one was much more

obviously comprehensive than the other (grouping items into threes was better than grouping them into twos). In the 200 list the common letters and the mirror-image groups yielded only groups of two items each, hence the vacillation.

In the 300 list there were very many competing groupings possible and there were no clear cut lines of demarcation among them (high complexity). Thus, while in the 100 and 200 lists the Ss could decide relatively easily on the structuring scheme, in the 300 list there were just too many alternatives available and none were entirely satisfactory. For example, VOF and VUK could be grouped together on the basis of the common 'V', or VOF could be grouped with QOS, XOQ, and SOZ on the basis of the common vowel. Alternately, QOS and QIJ could be grouped together on the basis of common first letter but QIJ could also be grouped with KUJ on the basis of a common last letter. Finally, some of the CVC's could also be grouped on the basis of pronunciation rather than on the basis of letter comunality. As can be seen the 300 list provides innumerable grouping alternatives, none of which are clear cut and mutually exclusive, nor is the relationship between the alternatives systematic. Consequently, the complexity of the list is very high. Furthermore, any attempt to form comprehensive groups leads to the formation of disjunctive rather than conjunctive groups and similarity is not transitive in disjunctive groups. For example, QIJ, QOS, and KUJ could be grouped together on the basis of the following grouping rule: 'common Q in the first position, or common J in the last position'. Such a rule, however, is not as satisfactory as a comprehensive conjunctive rule. In general, conjunctive groupings are preferred (Bruner, Goodnow, and Austin, 1956). Since the 300 list does not lend itself to the formation of comprehensive groupings and because

there is a multiplicity of primary grouping schemes available, the 300 list was learned much more slowly than the other lists. This may be taken as an alternate explanation of what was usually explained on the basis of interference and stimulus generalization. Such an explanation does not require a mechanistic conception of the Ss but explains interference while considering the Ss as cognitive beings.

In the 400 list there are very few groupings available and the list is basically quite different from both the 300, and the 100-200 type of list. In fact, it is not at all certain that all four lists are on the same continuum. Lists 100-200 and 400 may be regarded as being approximately at the vertices of a three dimensional space with performance as the vertical axis. As one goes from high similarity (OP based) to total difference it should be possible step-wise to reduce the similarity to nil but still to maintain its regular character, that is, to decrease OP in terms of groupability and comprehensiveness while maintaining complexity at a constant minimal level (as in 100-200, but not in 300). Alternately, it should be possible to start with the 300 list which is visualized as being at a different vertex from that of the 100-200 and 400 lists (see Figure 18), and to decrease complexity while maintaining the same level of comprehensiveness and groupability and to arrive at the 400 vertex. Finally, by changing all three it may be possible to go from 300 to 100-200. In other words, it is possible to view the 100-200 and 400 lists as approximately equal in complexity (generally low) and groupability (generally high), but varying in comprehensiveness. On the other hand, the 300 and 400 lists may be considered as approximately equal in groupability (generally high) and comprehensiveness (generally low), but differing in complexity.

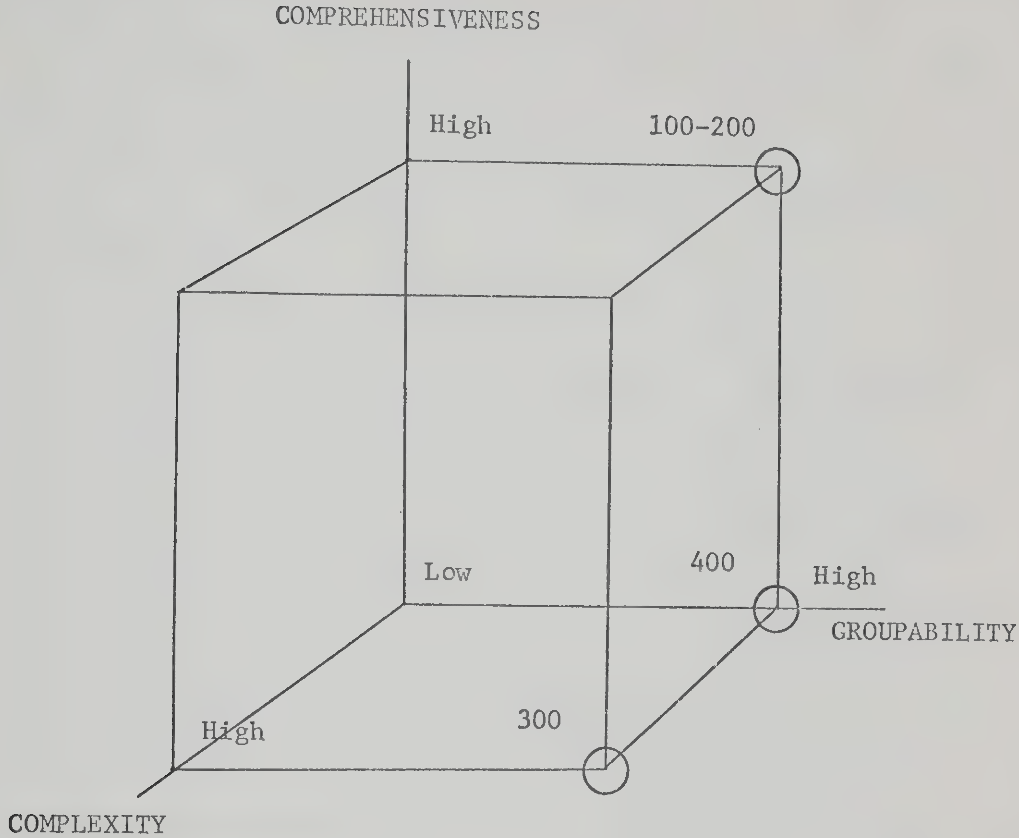


Figure 18. Diagrammatic Representation of the Lists in Terms of the Parameters of OP.

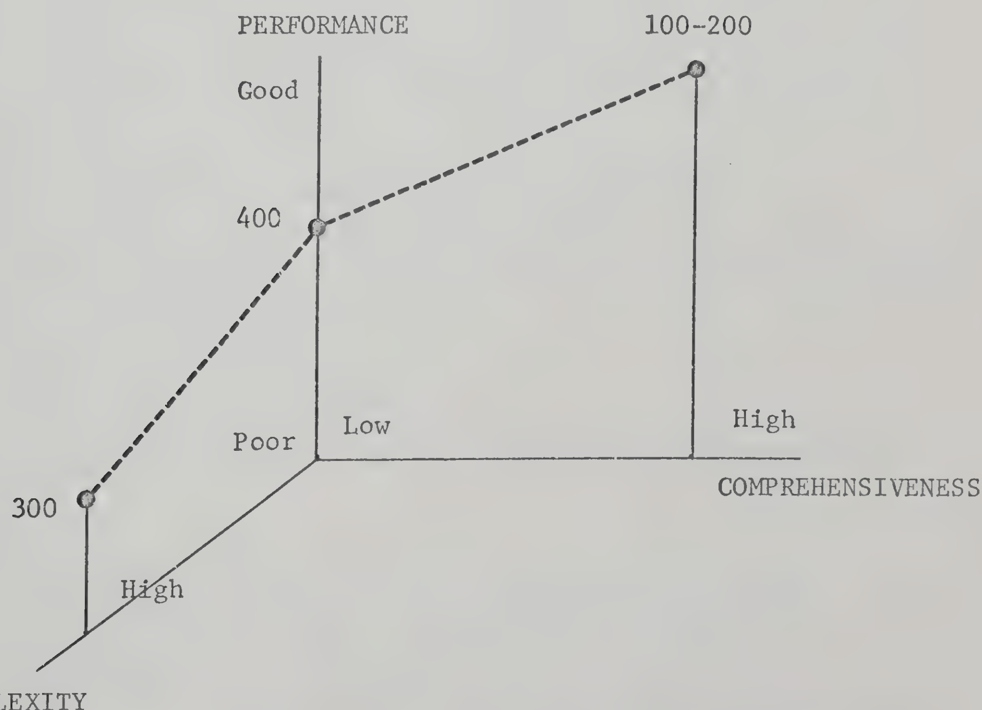


Figure 19. Performance on Experimental Lists in Relation to the Degree of Complexity and Comprehensiveness at a Common Level of Groupability.

The groupability in all cases is considered as high since it was possible to find a basis on which the items could be grouped into sets of three or four. (The square root of n, where n is the number of items, is regarded as being the number of groups yielding the highest degree of groupability).[1]

Consequently, there is no reason to expect performance to decrease from 100-200, to 300, to 400 (see Figure 19). The relationship to performance may be understood in terms of projections from three dimensions onto two by considering only the difference in performance between 300 and 400, or between 400 and 100-200. A comparison between 300 and 100-200 would involve variation in two parameters simultaneously, namely, comprehensiveness and complexity.

So far studies in verbal learning using nonsense StOs have explored the 300-400 continuum, while at times straying towards the 100-200 vertex. As a result they have obtained seemingly contradictory results since the 300-400 continuum and the 400-(100-200) continuum lie along different axes, and the data cannot be interpreted on the basis of one axis only.

In subsequent studies it should be profitable to keep this relationship in mind and therefore, to vary overall similarity by varying only one of the parameters of OP.

In general it may be predicted that performance will be better at high degrees of groupability, with high degrees of comprehensiveness

---

[1]In this interpretation the alphabetical basis for OP in the 400 list is not being differentiated from a similarity base since performance is being explained. Furthermore, there will always be some other base as a confounding factor in lists which attempt to be 'different' since in such circumstances any weak base becomes prominent.

and with low degrees of complexity (that is, at a high degree of similarity).

This statement, of course, assumes that the comparison among performances for different levels of similarity will be made at common levels of detail since StOs which require that organization occur at a low level of detail are not as easily learned as those which are organizable at a grosser level of detail. Thus, it is not very useful to compare the performance on meaningful words with that on nonsense words, except as an illustration that learning is easier at grosser levels of detail than at finer levels of detail. Similarly, lists which require Ss to organize on the basis of similarities among letter-shapes or letter-sounds are at a lower level of detail than lists which can be organized on the basis of common letters; that is, treating letters as the units of comparison rather than their individual shapes or sounds.

Therefore, if a common level of detail is assumed, performance should vary directly with similarity, keeping in mind the three parameter nature of OP in terms of which overall list similarity is defined.

In the present study it was also indicated that performance is a function of the learning conditions. The relationship between performance and the two variables of similarity and learning conditions may best be described in terms of pictorial representations. Figures 20 and 21 show the possible interactions. As the learning conditions become more free, permitting the Ss to organize the StOs, performance will improve for a given set of StOs. However, as similarity varies the performance will depend on the type of learning conditions selected. If these are permissive or free as in the present study, then

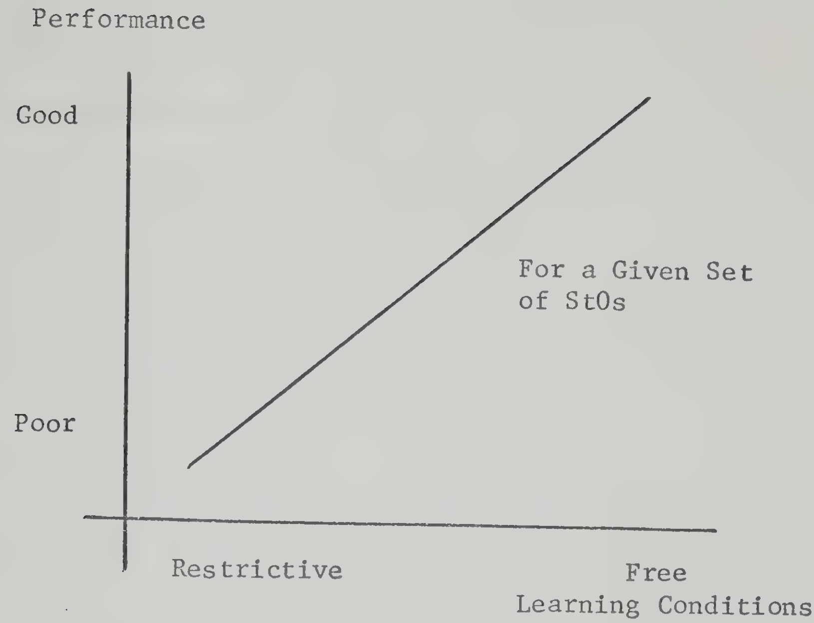


Figure 20. Relationship Between Performance and Learning Conditions for a Given Set of StOs.

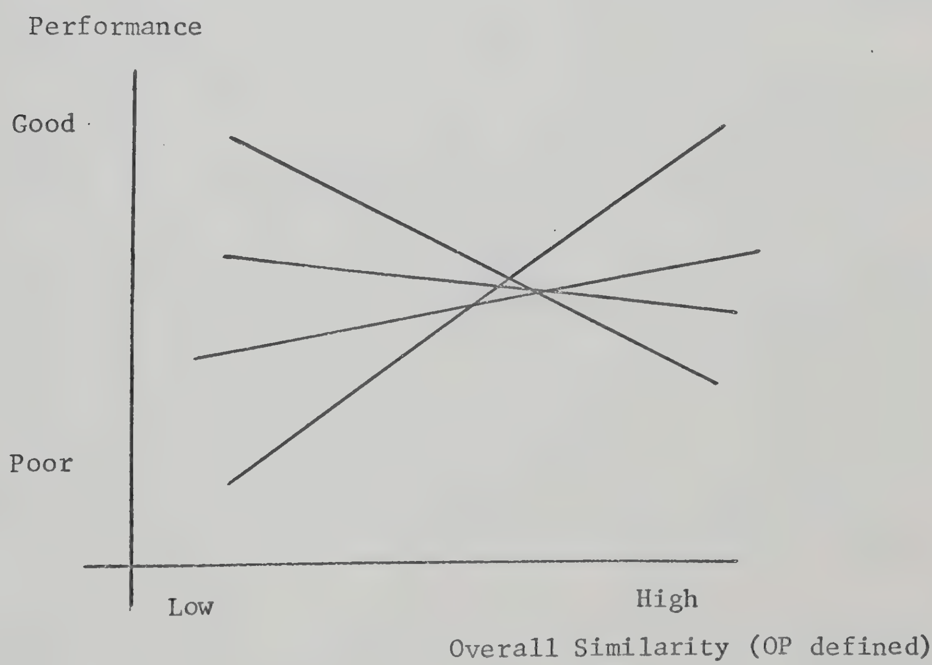


Figure 21. Relationship Between Performance and OP Defined Similarity Under Different Learning Conditions.

performance will improve with an increase in similarity; but if the learning conditions are highly restrictive, then an increase in similarity (OP based) may not necessarily lead to facilitation. The exact relationship between the variables can be determined experimentally in subsequent research. The determination of the pivotal point 'A' in Figure 21 should also be of theoretical interest since in that general range it would be possible to obtain equally good performance under different learning conditions.

The predicted relationship between learning conditions (free versus restrictive), however, must not be confused across different learning tasks, that is, across P-A, serial, and free recall. Although it is true that the latter is the freest in terms of opportunity to organize while the first is most restrictive, nevertheless, the learning tasks are quite different and thus any application of the predictions must be modified with respect to these differences.

Essentially the P-A, serial, and free recall tasks have a different purpose and involve different goals. In a P-A experiment the task of the Ss is, essentially, to learn new labels for a set of stimuli. This may be regarded as the very simplest type of concept learning (not concept attainment), namely, labelling. Any difficulty experienced by the Ss is not necessarily due to similarity *per se*, but to the learned habits developed since childhood in language learning. Generally, for example, a child learns to respond 'cat' to a large variety of relatively different StOs since 'cats' come in many different shapes and colours, ranging from house cats to jungle cats. Thus, in a P-A task the first problem encountered by the Ss is the need to work at a very fine level of detail, and the finer the level the more difficult the task. The

general set of the Ss is to classify StOs in terms of rather large classes. In the P-A task the problem is complicated since the Ss are only given one exemplar of each 'concept' for which they have to learn the label. Consequently, any similarity in the stimulus set will tend to force the Ss to organize at a lower level of detail. However, this depends on the nature of the similarity in the stimulus and response set. For example, similarity between a given stimulus and its corresponding response will facilitate learning; on the other hand, similarity among the stimuli may at times be used for grouping the stimuli together. Both Underwood (1961, p. 214) and Runquist (1966) have proposed this as a possible explanation for findings which are not quite consistent with interference theory. Runquist states:

> ...learning may be conceptualized as consisting of response learning, discrimination, association, and, in the case of high-similarity pairs of stimuli, the association of the particular responses which go with each pair of similar stimuli; i.e., S may learn that WIH and WIJ go with "wagon" and "insect" but not know which is which. If this latter phase were relatively complete by the time response learning had been accomplished, it might reduce the effects of list length by changing the task from one of learning 16 separate associations to simply learning eight discriminations (p. 11).

Consequently, any regularity which may be used by the Ss will tend to facilitate learning. Since similarity (OP based) provides a good basis for organization the presence of a high degree will tend to facilitate, and a low degree will tend to hinder learning as long as both are compared at equal levels of detail at which the Ss have to work.

In serial learning experiments the task is different in that the Ss have to learn not so much the items but the sequence that they are to form. This is akin to learning the alphabet or a telephone number. Thus, high similarity (OP based) should tend to facilitate

learning since it would provide patterns of regularity which the Ss could use to structure their recall. If the similarity is low (as in the 300 list) learning should be very difficult since there would be too many available relationships possible, none of which could be used to structure the list into a comprehensive whole. When a list possesses no similarity, other bases of OP may be used, the most common being alphabetical order, and performance would depend on the adequacy of this built in OP.

In a free learning experiment the learning task is the actual learning of the items and therefore it is most closely related to the learning of new things, not just new labels for old things. A free learning task is facilitated by a high degree of similarity (OP based) and retarded by low similarity per se since in the former the Ss may use the regularity provided by the similarity to structure the material into an integrated whole. In other words, highly similar material, because it can be structured, can be learned faster than material which is different.

This, of course, has significant applications to the school situation, particularly to the learning of science where similarity is such a ubiquitous phenomenon.

However, any direct application of experimental or 'laboratory' findings to the real class-room situation has to be made with great caution since the learning conditions and materials are rather different in each case. Nevertheless, an experimental study may serve the purpose of identifying variables that should be regarded as possibly significant in the applied type of learning experiments. Also, such research should be able to point out the trends or effects that might be expected.

Of course, the better understood a phenomenon is the more easily it is to apply to different conditions, and to predict more accurately its possible effects.

Similarity has been explicated to such an extent that it seems possible to predict what its effects might be on learning in a class-room setting. The study, first of all, indicated that similarity is ubiquitous and has potential for organization (Proposition II); second, it indicated that regardless of the meaninglessness of the task the Ss must be regarded as active strategy seeking organisms (Proposition I and the analysis of secondary data); third, the study showed that organization facilitates learning (Proposition III and examination of the order variable by experiment); and fourth, that a high degree of similarity (OP based) will facilitate learning.

On the basis of the present study, the variables that may be significant in class-room research are the following: the purposive behaviour of the Ss in seeking learning strategies, the impact of restrictive learning conditions which curtail such behavior, the importance of organizational potential of the material, and the facilitative effect of similarity in serving as a means of integrating ideas, concepts, processes, facts, etc. into comprehensive wholes.

It may be hypothesized that school children will learn faster under conditions which provide them with a maximum degree of freedom to structure what they are learning regardless of whether it is a set of facts or lawful relationships discovered in experiments. Furthermore, that children may be regarded as having a set to seek regularity and structure and to relate ideas into conjunctive hierarchical groups, rather than to learn these by rote without relating them in any way.

However, in order for them to do so the learning conditions must be right and similarity must be high rather than low.

This suggests that a useful task of applied research would be to determine the turning point at which similarity starts to facilitate learning by providing sufficient basis for organization. At low levels of similarity hindrance may be expected because of the complexity parameter of OP.

As a result of such applied research it should be possible to structure science learning experiences in such a way that it would maximize creative and inductive beharior on the part of the students. This is desirable since it would give the students some idea of the actual behavior of scientists and of their activity. As a result the students might be provided with a useful model of thinking and learning.

BIBLIOGRAPHY

Abramczyk, R.R., and Rousfield, W.A. Sequential ordering in repeated free recall as a function of interitem associative strength. Psychol. Rev., 1967, 17(2), 183-192.

Allen, M., Puff, C.R., and Weist, R. The effects of associative and coding processes on organization in free recall. J. Verb. Learn. verb. Behav., 1968, 7, 531-538.

Asch, S.E., Hay, J., and Diamond, R.M. Perceptual organization in serial rote-learning. Amer. J. Psychol., 1960, 73, 177-198.

Attneave, F.A. Method of graded dichotomies for the scaling of judgements. Psychol. Rev., 1949, 56,334-340.

Attneave, F. Dimensions of similarity. Amer. J. Psychol., 1950, 63, 516-555.

Baddeley, A.D. Language habits, S-R compatability and verbal learning. Amer. J. Psychol., 1964, 77, 463-468.

Baddeley, A.D. The influence of acoustic and semantic similarity on long term memory for word sequences. Quart.J.exp. Psychol., 1966, 18, 302-309.

Battig, W.F. Evidence for coding processes on 'rote' paired-associate learning. J. verb. Learn. verb. Behav., 1966, 5, 177-181.

Bay, K. Application of linear model simulator to educational and psychological research: A technical note on a computer program. Edmonton: Division of Educational Research Services, University of Alberta, 1970. (Mimeographed.)

Berlyne, D.E. Effects of spatial order and inter-item interval on recall of temporal order. Psychon. Sci., 1966, 6(8), 375-376.

Bourne, L.E., and Parker, B.K. Interitem relationships, list structures and verbal learning. Canad. J. Psychol., 1964, 18(1), 52-61.

Bousfield, A.K., and Bousfield, W.A. Measurement of clustering and of sequential constancies in repeated free recall. Psychol. Rep., 1966, 19(3pt.1), 935-942.

Bousfield, W.A. The occurrence of clustering in the recall of randomly arranged associates. J. gen. Psychol., 1953, 49, 229-240.

Bousfield, W.A., and Abramczyk, R.R. Sequential ordering in repeated free recall as a function of length of the stimulus word list. Canad. J. Psychol., 1966, 20, 427-434.

Bousfield, W.A., and Cohen, B.H. The effects of reinforcement on the occurrence of clustering in the recall of randomly arranged associates. J. Psychol., 1953, 36, 67-81.

Bousfield, W.A., Cohen, B.H., and Whitmarsh, G.A. Associative clustering in the recall of words of different taxonomic frequencies of occurrence. Psychol. Rep., 1958, 4, 39-44.

Bousfield, W.A., and Puff, C.R. Clustering as a function of response dominance. J. exp. Psychol., 1964, 67, 76-79.

Bousfield, W.A., Puff, C.R., and Cowan, T.M. The development of constancies in sequential organization during repeated free recall. J. verb. Learn. verb. Behav., 1964, 3, 489-495.

Brown, R.W. Language and Categories. In J.S. Bruner, J.J. Goodnow, and G.A. Austin. A Study of Thinking. New York: Wiley, 1956.

Brown, S.C., and Read, J.D. Effect of simultaneous list presentation on paired-associate learning. Psychol. Rep., 1966, 19, 787-790.

Bruner, J.S., Goodnow, J.J., and Austin, G.A. A Study of Thinking. New York: Wiley, 1956.

Carluccio, L., and Crowder, R.G. Constant order of pairs in the presentation and testing of paired associates. J. exp. Psychol., 1966, 72, 614-616.

Carnap, R. Logical Foundations of Probability. Chicago: The University of Chicago Press, 1950.

Carterette, E.C. A replication of free recall and ordering of trigrams. J. exp. Psychol., 1963, 66, 311-313.

Chambliss, D. The relation between judged similarity and the physical properties of plane figures. Unpublished Doctor's dissertation, University of Wisconsin, 1957.

Chomsky, N., and Miller, G.A. Introduction to the formal analysis of natural languages. In P.D. Luce, R.R. Bush, and E. Galanter (Eds), Handbook of Mathematical Psychology. New York: Wiley, 1963.

Cofer, C.N. A study of clustering in free recall based on synonyms. J. gen. Psychol., 1959, 60, 3-10.

Cofer, C.N., and Bruce, D.R. Clustering in free recall as a function of certain methodological variations. J. exp. Psychol., 1966 71, 858-866.

Cohen, B.H. An investigation of recoding in free recall. J. exp. Psychol., 1963, 65, 368-376.

Dallett, K.M. Number of categories and category information in free recall. J. exp. Psychol., 1964, 68, 1-12.

Deese, J. On the structure of associative meaning. Psychol. Rev., 1962, 69, 161-175.

Eagle, M., and Leiter, E. Recall and recognition in intentional and incidental learning. J. exp. Psychol., 1964, 68, 58-63.

Earhard, M. Subjective organization and list organization as determinants of free recall and serial recall memorization. J. verb. Learn. verb. Behav., 1967, 6(4), 501-507.

Eckstrand, G.A., and Wickens, D.D. Transfer of perceptual set. J. exp. Psychol., 1954, 47, 274-278.

Edwards, A.L. Experimental Design in Psychological Research. New York: Rinehart and Co. Inc., 1960.

Ekstrand, B., and Underwood, B.J. Paced versus unspaced recall in free learning. J. verb. Learn. Behav., 1963, 2, 288-290.

Epstein, W. The influence of syntactical structure on learning. Amer. J. Psychol., 1961, 74, 80-85.

Faber, I.E. The things people say to themselves. Amer. Psychologist, 1963, 18, 185-197.

Feldman, S.M., nad Underwood, B.J. Stimulus recall following paired-associate learning. J. exp. Psychol., 1957, 53, 11-15.

Forgus, R.H. Perception. New York: McGraw-Hill, 1966.

Forster, K.I. The effect of syntactic structure on nonordered recall. J. verb. Learn. verb. Behav., 1966, 5(3), 292-297.

Frincke, G. Word characteristics, associative-relatedness, and the free recall of nouns. J. verb. Learn. verb. Behav., 1968, 7, 366-372.

Gagne, R.M. The effect of sequence of presentation of similar items on the learning of paired associates. J. exp. Psychol., 1950, 40, 61-73.

Garner, W.R. Uncertainty and Structure as Psychological Concepts. New York: Wiley, 1962.

Garner, W.R., and Whitman, J.R. Form and amount of internal structure as factors in free-recall learning of nonsense words. J. verb. Learn. verb. Behav., 1965, 4, 257-266.

Gibson, E.J. The formation and discovery of 'high-order' units in intellectual tasks. In J.S. Bruner (Ed.), Learning About Learning. Washington: U.S. Government Printing Office, 1966.

Glanzer, M., and Meinzer, A. The effects of intralist activity on free recall. J. verb. Learn. verb. Behav., 1967, 6, 928-935.

Goldmeier, E. Uber Ahnlichkeit bei gesehenen Figuren. Psychol. Forsch., 1937, 21, 146-208.

Gulliksen, H. Paired comparisons and the logic of measurement. Psychol. Rev., 1946, 53, 189-213.

Haagen, C.H. Synonymity, vividness, familiarity and association value ratings of 400 pairs of common adjectives. J. of Psychol., 1949, 27, 453-463.

Heidbreder, E. The attainment of concepts. III. The process. J. Psychol., 1947, 24, 93-138.

Homzir, M.J., and Weimer, J. Connotative similarity and paired-associate learning. J. verb. Learn. verb. Behav., 1967, 6, 136-138.

Horowitz, L.M. Free recall and ordering of trigrams. J. exp. Psychol., 1961, 62, 51-57.

Horowitz, L.M. Associative matching and intralist similarity. Psychol. Rep., 1962, 10, 751-757.

Hull, C.L. Quantitative aspects of the evolution of concepts: An experimental study. Psychol. Monogr., 1920, 28 123.

Hume, D. An Enquiry Concerning Human Understanding. Vol. 35 in R.M. Hutchins (Ed.), Great Books of the Western World. Chicago: Enc. Brit. Inc., 1952, pp. 451-509.

James, W. The Principles of Psychology. Vol. 53 in R.M. Hutchins (Ed.), Great Books of the Western World. Chicago: Enc. Brit. Inc., 1952.

Jenkins, H.M. Formation of concepts under noncontingent reinforcements. Paper presented at Annual meeting of AAAS, New York, 1960. Reported in Shepard, R.N. Comments on Professor Underwood's Paper. In C.N. Cofer and B.S. Musgrave (Eds.), Verbal Behaviour and Learning. New York: McGraw-Hill, 1963, pp. 61-62.

Jenkins, J.J., Mink, W.D., and Russell, W.A. Associative clustering as a function of verbal association strength. Psychol. Rep., 1958, 4, 127-136.

Jenkins, J.J., and Russell, W.A. Associative clustering during recall. J. Abnorm. soc. Psychol., 1952, 47, 818-821.

Jensen, A.R., and Rohwer, W.D. Jr. The effect of verbal mediation on the learning and retention of paired-associates retarded adults. Amer. J. ment. Defic., 1963(a), 68, 80-84.

Jensen, A.R., and Rohwer, W.D. Jr. Verbal mediation in paired associate and serial learning. J. verb. Learn. verb. Behav., 1963(b) 1, 346-352.

Karwoski, T. The memory value of size. J. exp. Psychol., 1931, 14, 539-554.

Karwoski, T., and Schacter, J. Psychological studies in semantics: III Reaction times for similarity and difference. J. soc. Psychol., 1948, 28, 103-120.

Katona, G. Organizing and Memorizing. New York: Columbia University Press, 1940.

Kausler, D.H. (Ed.) Readings in Verbal Learning. New York: Wiley, 1966.

Keenan, V.C. Presentation method and learning - Time constancy in serial learning. Unpublished Doctor's dissertation, University of California, 1967. Dissertation Abstract.

Koeppel, J.C. Intralist similarity, internal structure, and free recall. Unpublished Doctor's dissertation, Oklahoma State University, 1967.

Koepel, J.C. Intralist similarity, internal structure, and free recall. J. verb. Learn. verb. Behav., 1968, 7, 882-886.

Koffka, K. Principles of Gesalt Psychology. New York: Harcourt and Brace, 1935.

Kluckholen, G., and Kelly, W.H. The concept of culture. In R. Linton (Ed.), The Science of Man in the World Crisis. New York: Columbia University Press, 1946.

Kuennapas, T. Visual memory of capital letters: multidimensional ratio scaling and multidimensional similarity. Percept. mot. Skills, 1967, 25, 345-350.

Landahl, H.D. Neural mechanisms for the concepts of difference and similarity. Buff. Math. Biophysics, 1945, 7, 83-88.

Langford, C.H. The notion of analysis in Moore's philosophy. In P.A. Schilpp (Ed.), The Philosophy of G.E. Moore. New York: Tudor Pub. Co., 1952.

Lawrence, D.H. The nature of the stimulus: Some relationships between learning and perception. In S. Koch (Ed.), Psychology: A Study of a Science. Vol 5, New York: McGraw-Hill, 1963.

Leicht, K.L., and Kausler, D.H. Supplementary report: Functional stimulus learning as related to degree of practice and meaningfulness. J. exp. Psychol., 1965, 69, 100-101.

Levitt, H., and Goss, A.E. Stimulus attributes and drive in paired-associate learning. J. exp. Psychol., 1961, 62, 243-252.

Lieberman, L.R., and Culpepper, J.T. Words versus objects: comparison of free verbal recall. Psychol. Rep., 1965, 17, 983-988.

Mandler, G. Verbal learning. In G. Mandler, G.P. Mussen, N. Kogan, and M.A. Wallach, New Directions in Pyschology III. New York: Holt, Reinhart and Winston, Inc., 1967.

Manis, M. Cognitive Processes. Belmot, California: Wadsworth Publishing Company, 1966.

Marks, L.E., and Miller, G.A. The role of the semantic and syntactic constraints in the morization of English sentences. J. verb. Learn. verb. Behav., 1964, 3, 1-5.

Marshall, G.R. Stimulus characteristics contributing to organization in free recall. J. verb. Learn. verb. Behav., 1967, 6, 354-374.

Martin, E. Stimulus recognition in aural paired - associate learning. J. verb. Learn. verb. Behav., 1967, 6(2), 272-276.

Martin, E. Recognition and correct responding mediated by first letter trigram stimului. J. verb. Learn. verb. Behav., 1968, 7(3), 703-706.

Martin, C.J., Boersma, F.J., and Cox, D.L. A classification of associative strategies in paired-associate learning. Psychon. Sci., 1965, 3, 455-456.

Martin, C.J., Cox, D.L., and Boersma, F.J. The role of associative strategies in the aquisition of P-A material: An alternative approach to meaningfulness. Psychon. Sci., 1965, 3, 463-464.

Martin, C.J., and Saltz, E. Serial versus random presentation of paired-associates. J. exp. Psychol., 1963, 65, 609-615.

McCullough, P. The perceptual discrimination of similarities. Unpublished Doctor's dissertation, University of Utah, 1957.

McGeoch, J.A., and Irion, A.L. The Psychology of Human Learning. New York: David McKay Company, Inc., 1952.

McGeoch, J.A., and McKinney, F. Studies in retrocative inhibition VIII: The influence of the relative order of presentation of original and interpolated paired-associates. J. exp. Psychol., 1937, 20, 60-83.

McGeoch, J.A., and Underwood, B.J. Tests of the two-factor theory of retroactive inhibition. J. exp. Psychol., 1943, 32, 1-16.

McCourty, M.C. Serial position effects on learning as a function of interfering associations. Unpublished Master's thesis, State University of Iowa, 1940.

McNulty, J. The measurement of 'adopted chunks' in free recall learning. Psychon. Sci., 1966, 4, 71-72.

Mellgren, R.L. Position as a stimulus in verbal learning. Psychon. Sci., 1967, 8(2), 73-74.

Messick, S.J., and Abelson, R.P. The additive constant problem in multidimensional scaling. Psychometrika, 1956, 21, 1-17.

Miller, G.A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychol. Rev., 1956, 63, 81-97.

Miller, G.A. Free recall of redundant srings of letters. J. exp. Psychol., 1958, 57, 484-491.

Miller, G.A., Bruner, J.A., and Postman, L. Familiarity of letter sequences and tachistoscopic identification. J. genr. Psychol., 1954, 50, 129-139.

Miller, G.A., Galanter, E., and Pribram, K. H. Plans and the Structure of Behavior. New York: Rinehart and Winston, Inc., 1960.

Miller, G.A., and Selfridge, J.A. Verbal context and the recall of meaningful material. Amer. J. Psychol., 1950, 63, 176-185.

Monty, R.A., Fisher, D.F., and Karsh, R. Stimulus characteristics and spatial encoding in sequential short-term memory. J. Psychol., 1967, 65, 109-116.

Moursund, J.P., and Chape, C.M. Comparison of three presentation methods in paired-associate learning. Psychon. Sci., 1966, 6(4), 167-168.

Muir, W.R. Measurement of similarity in verbal material by a paired-comparison procedure. Unpublished Doctor's dissertation, University of Alberta, 1968.

Newman, B.B. Forgetting of meaningful material during a sleep and waking. Amer. J. Psychol., 1939, 52, 65-71.

Newman, S. Serial position as a cue in learning: the effect of test rate. J. exp. Psychol., 1966, 71, 319-320.

Newman, S.E. Multitrial free recall: effects of clustered presentation, element-sharing, and instructions. J. verb. Learn. verb. Behav., 1967, 6, 844-850.

Newman, S.E., and Buckhout, R. S-R and R-S learning as functions of intralist similarity. Amer. J. Psychol., 1962, 75, 429-436.

Newman, S.E., and Saltz, E. Serial position as a cue in learning Amer. J. Psychol., 1962, 75, 102-108.

Noble, C.E. Psychology and the logic of similarity. J. gen. Psychol., 1957, 57, 23-43.

Nodine, C.F., and Hardt, J.V. Generated meaningfulness of single letters as a function of position in CVC trigrams. Psychon. Sci., 1968, 10(4), 129-130.

Optical Soc. Amer., Committee on colorimetry: 1953. The Science of Color. New York: Cromwell.

Osgood, C.E. Method and Theory in Experimental Psychology. New York: Oxford University Press, 1953.

Osgood, C.E. Studies in the generality of affective meaning systems. Amer. Psychologict, 1962, 17, 10-28.

Plato, Parmenides. Vol. 7 in R.M. Hutchins (Ed.), Great Books of the Western World. Chicago: Enc. Brit. Inc., 1952. pp. 486-511.

Pollio, H.R. Composition of associative clusters. J. exp. Psychol., 1964, 67, 199-208.

Postman, L., and Egan, J.E. Experimental Psychology: An Introduction. New York: Harper and Brothers, 1949.

Postman, L., and Riley, D.A. A critique of Kohler's theory of association. Psychol. Rev., 1957, 64, 61-72.

Postman, L., and Senders, V.L. Incidental learning and generality of set. J. exp. Psychol., 1946, 36, 153-165.

Puff, C.R., and Bousfield, W.A. Development of intertrial organization with simultaneous presentation of stimulus items. J. verb. Learn. verb. Behav., 1967, 6, 213-215.

Richardson, M.W. Multidimensional psychophysics. Psychol. Bull., 1938, 35, 659-660.

Rotberg, I.C., and Woolman, M. Verbal paired-associate learning as a function of grouping similar stimuli or responses. J. exp. Psychol., 1963, 65, 47-51.

Rothkopt, E.Z., and Coke, E.U. The prediction of free recall from word association measures. J. exp. Psychol., 1961, 62, 433-438.

Rubin, E.D., and Brown, S.C. Constant versus varied serial order in paired-associate learning: The effect of formal intralist similarity. J. exp. Psychol., 1967, 73, 257-262.

Runquist, W.N. Intralist interference as a function of list length and interstimulus similarity. J. verb. Learn, verb. Behav., 1966, 5, 7-13.

Runquist, W.N. Formal intralist stimulus similarity in paired-associate learning. J. exp. Psychol., 1968(a), 78, 634-641.

Runquist, W.N. Rated similarity of high m CVC trigrams and Words and low m CCC trigrams. J. verb. Learn. verb. Behav., 1968(b), 7, 967-968.

Runquist, W.N. Structural effects of letter identity among stimuli in paired-associate learning. J. exp. Psychol., 1970(a), 84, 152-163.

Runquist, W.N. Acoustic similarity among stimuli as a source of interference in paired-associate learning. J. exp. Psychol., 1970(b), 83, 319-322.

Runquist, W.N., and Joinson, P.A. Rated similarity of low association value CVC trigrams. J. verb. Learn. verb. Behav., 1968, 7, 317-320.

Samuels, S.J., and Jeffrey, W.E. Locus of facilitation for paired-associate learning under fixed order presentation. Psychon. Sci., 1968, 10(10), 355-356.

Segal, E.M., and Halwes, T.G. Learning of letter pairs as a prototype of first language learning. Psychon. Sci., 1965, 3, 451-452.

Segal, E.M., and Halwes, T.G. The influence of frequency of exposure on the learning of phrase structural grammar. Psychon. Sci., 1966, 4, 157-158.

Senter, R.J., and Hauser, G.K. An experimental study of a mnemonic system. Psychon. Sci., 1968, 10(8), 289-290.

Shepard, R.N. The analysis of proximities: Multidimensional scaling with unknown distance function I. Psychometrika, 1962(a), 27, 125-140.

Shepard, R.N. The analysis of proximities: Multidimensional scaling with unknown distance function II. Psychometrika, 1962(b), 27, 219-246.

Shepard, R.N. Comments on Professor Underwood's paper. In C.N.Cofer, and B.S. Musgrave (Eds.), Verbal Behavior and Learning. New York: McGraw-Hill, 1963.

Shepard, R.N. Attention and the metric structure of the stimulus space. J. math. Psychol., 1964, 1, 54-87.

Shepard, R.N. Learning and recall as organization and search. J. verb. Learn. verb. Behav., 1966, 5, 201-204.

Shepard, R.N., Hovland, C.I., and Jenkins, H.M. Learning and memorization of classifications. Psychol. Monogr., 1961, 75, 13.

Smith, K.H. Mediation and position learning in the recall of structured letter pairs. Psychon. Sci., 1965, 2, 293-294.

Spiker, C.C. Associative transfer in verbal paired-associate learning. Child Develpm., 1960, 31, 73-87.

Sundland, D.M., and Wickens, D.D. Context factors in paired associate learning and recall. J. exp. Psychol., 1962, 63, 302-306.

Torgerson, W.S. Multidimensional scaling: I. Theory and Method. Psychometrika, 1952, 17, 401-419.

Torgerson, W.S. Theory and Methods of Scaling. New York: Wiley, 1958.

Torgerson, W.S. Multidimensional scaling of similarity. Psychometrika, 1965, 30, 379-393.

Tulving, E. Subjective organization in free recall of 'unrelated' words. Psychol. Rev., 1962(a), 69, 344-354.

Tulving, E. The effect of alphabetical subjective organization on memorizing unrelated words. Canad. J. Psychol., 1962(b), 16, 185-191.

Umemoto, T. Paired-associate learning as a function of similarity: Semantic similarity between stimulus and response-items. Amer. J. Psychol., 1962, 75, 85-93.

Umemoto, T., and Hilgard, E.R. Paired-associate learning as a function of similarity: Common stimulus and response items within the list. J. exp. Psychol., 1961, 62, 97-104.

Underwood, B.J. Studies in distributed practice. VII. Learning and retention of serial nonsense lists as a function of intralist similarity. J. exp. Psychol., 1952, 44, 80-87.

Underwood, B.J. Studies in distributed practice: VIII. Learning and retention of paired nonsense syllables as a function of intralist similarity. J. exp. Psychol., 1952(a), 45, 133-141.

Underwood, B.J. Studies in distributed practice: IX. Learning and retention of paired adjectives as a function of intralist similarity. J. exp. Psychol., 1953(b), 4, 143-149.

Underwood, B.J. Studies in distributed practice: X. The influence of intralist similarity on learning and retention of serial adjective lists. J. exp. Psychol., 1953(c), 45(4), 253-259.

Underwood, B.J. An evaluation of the Gibson Theory of verbal learning. In C.N. Cofer (Ed.), Verbal Learning and Verbal Behavior. New York: McGraw-Hill, 1961.

Underwood, B.J. Stimulus selection in verbal learning. In C. N. Cofer, and B.S. Musgrave (Eds.), Verbal Behavior and Learning. New York: McGraw-Hill, 1963.

Underwood, B.J. The representativeness of rote verbal learning. In A.W. Melton (Ed.), Categories of Human Learning. New York: Academic Press, 1964.

Underwood, B.J. Experimental Psychology. (Second Edition) New York: Division of Meredith Publishing Co., 1966.

Underwood, B.J., Ekstrand, B.R., and Keppel, G. Studies in distributed practice: XXIII. Variations in response-term interference. J. exp. Psychol., 1964, 68, 201-212.

Underwood, B.J., and Goad, B. Studies in distributed practice: I. The influence of intralist similarity in serial learning. J. exp. Psychol., 1951, 42, 125-134.

Underwood, B.J., Ham, M., and Ekstrand, B. Cue selection in paired-associate learning. J. exp. Psychol., 1962, 64, 405-409.

Underwood, B.J., and Richardson, J. Studies in distributed practice: XVII. Interlist interference and the retention of paired consonant syllables. J. exp. Psychol., 1957, 54, 274-279.

Underwood, B.J., and Richardson, J. Studies in distributed practice: XVIII. The influence of meaningfulness and intralist similarity of serial nonsense lists. J. exp. Psychol., 1958, 56(3), 213-219.

Underwood, B.J., Runquist, W.N., and Schulz, R.W. Response learning in paired-associate lists as a function of intralist similarity. J. exp. Psychol., 1959, 58(1), 70-78.

Underwood, B.J., and Schulz, R.W. Meaningfulness and Verbal Learning. Chicago: J.B.Lippincott Co., 1960.

Wallach, M. A. On psychological similarity. Psychol. Rev., 1958, 65(2), 103-116.

Wallace, W.H., Turner, S.H., and Perkins, C.C. Preliminary studies of human information storage. Project No. 132C, 1957, Institute for Cooperative Research, University of Pennsylvanie. Reported in G.A. Miller, E. Galanter, and K. H. Pribram, Plans and the Structure of Behavior. New York: Rinehart and Winston, Inc., 1960.

Whitman, J.R. Form of internal and external structure as factors in free-recall and ordered recall of nonsense and meaningful words. J. verb. Learn. verb. Behav., 1966, 5, 68-74.

Wickens, D.D. Conference Summary. In C.N. Cofer, and B.S. Musgrave (Eds.), Verbal Behavior and Learning. New York: McGraw-Hill, 1963.

Wiener, B.J. Statistical Principles in Experimental Design. New York: McGraw-Hill Books Co., 1962.

Woodworth, R.A., and Schlosberg, H. Experimental Psychology. (Revised Edition) New York: Holt, 1954.

Young, G., and Householder, A.S. Discussion of a set of points in terms of their mutual distances. Psychometrika, 1938, 3, 19-22.

APPENDIX A

Definitions and Abbreviations

DEFINITIONS AND ABBREVIATIONS

CCC - a trigram (nonsense syllable) consisting of three consonants (C), e.g. GXT.

COMPLEXITY - the degree and quality of the multiplicity of available alternatives for the formation of primary and secondary groups in a list (see pp. 123-124).

COMPREHENSIVENESS - the degree to which all the items in a list can be encompassed in one simple coding statement (see pp. 121-123).

CONTEXT EFFECT - the influence that all the items in a list have on the choice of coding mode (spelling versus pronunciation) that will be used in learning a given item in that list (see pp. 171-179).

CVC - a trigram (nonsense syllable) consisting of a consonant (C), vowel (V), and a consonant (C), e.g. VOF.

_E_ - the experimenter

ENVIRONMENTAL CHARACTERISTIC - a property that relates the StO to the environment. For example, an environmental characteristic of the StO, table, is that it is located in a dining room (see pp. 35-36).

FORMAL SIMILARITY - the number of common letters (elements) per list (see p. 25, and p. 120).

GROUPABILITY - the potential of grouping the items in a list into more than one group of all _n_ items, and into less than _n_ groups of one item. Maximum groupability is n, where _n_ is the number of items in the list (see p. 121)

HfS - high formal similarity

HOP - high organizational potential

INTERNAL ORGANIZATION - the sequential constraints and spatial coordinates existing among the elements of a StO (see p. 110).

INTRINSIC CHARACTERISTIC - a property displayed by the StO itself (see pp. 35-36).

LEVEL OF DETAIL - the degree of refinement in a comparison. For example, at a gross level of detail a tiger and a house cat may be regarded as similar; at a fine level of detail an angora cat and a siamese cat may be regarded as different (see pp. 48-50).

LfS - low formal similarity

LOP - low organizational potential

OP - organizational potential

ORGANIZATIONAL POTENTIAL - a potential for organization is said to exist in a list of the items if the list can be grouped (organized) into a pattern or into classes (see pp. 64-72, and pp. 121-124).

OVERALL LIST SIMILARITY - the degree of OP-defined similarity in a list as a whole (see also similarity, pp. 114-118, and p. 121).

OVERALL SIMILARITY - see overall list similarity

PAIR-WISE SIMILARITY - the similarity between two items in a list with the degree of similarity being determined by the number of common elements (letters) in common positions with common sequential constraints (see pp. 114-118, and pp. 120-121).

R - (as in 100R, 200R, 300R, and 400R) - random list - a list designated as R is one in which the item order (sequence) was determined on the basis of a table of random numbers (see pp. 125-130).

S - (as in 100S, 200S, 300S, and 400S) - ordered list - a list designated as S is one in which the item order (sequence) selected emphasized a built in organizational potential (see pp. 125-130).

<u>S</u>, <u>Ss</u> - subject, subjects - the learner(s) in a verbal learning experiment.

SEQUENTIAL CONSTRAINTS - the relationship between a given element in a StO and the following element in that StO (see pp. 107-109).

SIMILARITY - similarity defined in terms of organizational potential, and therefore, variation in the degree of similarity would be directly related to the variation in the degree of organizational potential (see Chapter 5, and pp. 120-124).

SPATIAL COORDINATES - the location of the elements in a given StO relative to some frame of reference (see also spatial relationships and pp. 107-109).

SPATIAL RELATIONSHIPS - the pattern formed by the common elements in a StO (see also spatial coordinates, pp. 41-43, and pp. 107-109).

StO, StOs - stimulus object(s)

STIMULUS OBJECT - any physical object, word, or abstract idea, etc. The symbol, StO, is parallel to x in mathematics.

TRANSITIVE - (as in a transitive relation) - a relation which has the property that if A bears the relation to B and B bears the same relation to C, then A bears the relation to C (see p. 44).

TRANSITIVITY - see transitive

W-METHOD - see whole method of presentation

W-PRESENTATION - see whole method pf presentation

WHOLE METHOD OF PRESENTATION - a simultaneous presentation of all the items in a list (see pp. 83-86).

APPENDIX B

Supervisor's Introductory Remarks to the Subjects

Supervisor's Introductory Remarks to the Subjects

Prior to the Learning Experiment the Supervisor made a number of general remarks which may be summarized as follows:

1. This is a learning experiment and not a test.

2. No special preparation or knowledge is needed.

3. It is a very important experiment and many such are conducted at universities and other high schools.

4. The procedure may seem strange, but there are many good reasons for it; and it is hoped that the experience proves to be interesting.

5. Any other information as to the experiment will be given at the conclusion of the session to those who are interested.

6. After the results have been analysed a summary will be sent to all teachers and interested students may in this way get more details.

APPENDIX C

General Instructions

This is a specimen set of the instructions handed out to the subjects and read to them prior to the learning task.

INSTRUCTIONS

This experiment is in two parts:

Part 1: Study and Recall

Part 2: Questionnaire

PART 1: Study and Recall

The object of Part 1 is to learn a list of 12 nonsense words as fast as possible. For this reason all of you will be asked to use a rather special learning method which is described below.

The nonsense words are also rather special. They all have 3 letters, and in every case the first and last letter is a consonant while the middle letter is a vowel. For example, KIV, GUP, DIV, etc.

Learning Method

I will say "Ready?"

"STUDY"

When I say "STUDY" you will turn over the Word List and start learning the words. (You will have exactly 2 minutes).

"TIME"

When I say "TIME" you will turn the Word List face down and attempt to write out in the Practice Booklet as many of the words as you can remember. YOU MAY WRITE THE WORDS IN ANY ORDER. Before you start writing check to see that the Back-up-Sheet is directly under the page on which you are about to write. (You will have exactly 1 minute).

"TIME"

When I say "TIME" you will stop writing and turn to the next blank page, folding the used one completely under the Practice Booklet. At this time you will also shift the Back-up-Sheet so that it is directly under the next blank page.

Then I will say: "Ready?"

"STUDY"

When I say "STUDY" you will go through the same procedure as above.

The whole cycle of studying and writing will be repeated five times by ALL students. After you have completed your fifth writing I will say:

"TIME, CHECK YOUR WORK"

If you have all 12 correct you will go on to Part 2 (Questionnaire), if not you will continue for five more "STUDY-WRITE" turns.

After you have completed your tenth writing I will again say:

"TIME, CHECK YOUR WORK"

If you have all 12 correct you may go on to Part 2 (Questionnaire), if not you will continue for five more "STUDY-WRITE" turns.

After the fifteenth writing ALL students will go on to Part 2, even those that still do not have all 12 correct. It simply means that you were unlucky and got one of the very difficult Word Lists. All lists vary in difficulty.

IT IS VERY IMPORTANT THAT YOU STOP ONLY AFTER 5, 10, or 15 WRITINGS AND NOT ANYWHERE ELSE even if you are absolutely sure that you have a perfect score. REMEMBER DO NOT STOP EXCEPT AFTER THE 5th, 10th, or 15th writing.

ARE THERE ANY QUESTIONS?

Example

Now let us practice the "STUDY-WRITE" Learning Method using the following four numbers:

492
638
704
498

Please use pages A, B, and C in the Practice Booklet for the writing part of the cycle. At this time your Practice Booklet should be open to page A and the Back-up-sheet should be directly under page A.

REMEMBER: YOU MAY WRITE DOWN THE ITEMS IN ANY ORDER.

APPENDIX D

Practice Booklet Specifications

and Title Page

Practice Booklet Specifications

Dimensions:	8 1/2 inches by 11 inches
Contents:	Title Page (Exemplar follows)
	18 blank pages
	one loose sheet called a Back-up-sheet
Paper:	Rapidrun 40M from the E.B. Eddy Company

The instructions for the use of the Back-up-sheet appeared on the Title Page of the Practice Booklet. Its purpose was to ensure that even when a S pressed very hard on one page the image would not pass through to the next page and interfere with unaided recall.

Code Name of Word List: ________

PRACTICE BOOKLET

1. As soon as you get this Booklet please take out the loose sheet of paper and label the pages inside as follows:

   Use the letters A, B, and C for the first three pages, and the numbers 1 to 15 for the rest.

   We will use pages A, B, and C for the example, and pages 1 to 15 for the actual learning experiment.

2. The loose sheet of paper is your Back-up-sheet.

   It should ALWAYS be directly under the page on which you are about to write.

APPENDIX E

Experimental Lists: A Specimen Set

VUK

VUT

VUX

KUV

TUV

XUV

ZOH

ZOJ

ZOS

HOZ

JOZ

SOZ

JOZ

ZOH

HOZ

VUK

SOZ

TUV

VUT

XUV

ZOJ

KUV

ZOS

VUX

QIF

QIJ

FIQ

JIQ

VUK

VUX

KUV

XUV

ZOH

ZOS

HOZ

SOZ

HOZ

KUV

ZOS

QIF

SOZ

VUK

QIJ

VUX

XUV

JIQ

ZOH

FIQ

FUX

HUZ

JIH

KUJ

QIJ

QOS

SOZ

VOF

VUK

XIF

XOQ

ZIV

XOQ

SOZ

XIF

FUX

ZIV

QIJ

HUZ

QOS

VOF

KUJ

VUK

JIH

BOF

CEH

GYX

LAJ

MYP

NID

QIJ

QUR

TEH

VAW

VUK

ZOS

VUK

QIJ

VAW

BOF

ZOS

MYP

CEH

NID

QUR

LAJ

TEH

GYX

APPENDIX F

The Questionnaire

<u>PART 2: QUESTIONNAIRE</u>

1. Code Name of Word List ______________

2. Your age __________ years

3. Grade __________

4. Sex: Male __________

   Female __________

Please answer the following questions clearly, accurately, and in detail.
Please write LEGIBLY.

1. In learning your list did you notice any similarities among the words? Yes or no? __________ If yes, please describe in detail what you noticed, give examples, and explain how you used it to help you to remember the words.

2. If you answered 'yes' to question 1, please state WHEN you first noticed the similarities. Use the page number of the Practice Booklet to pinpoint the time.

3. Did you group any of the words into pairs or larger groups? Yes or no? __________ If yes, please describe exactly and give examples.

4. Did you use alphabetical order to help you to remember the words? Yes or no? __________ If yes, explain how useful this was and if you used it for all 12 words. If you did not use it for all 12, for which ones <u>did</u> you use alphabetical order?

5. Did you remember some of the words because they were similar to real words? Yes or no? __________ If yes, please list the nonsense words and the real words which they resemble.

6. Did you try to make up a sentence using some of the words or the first letters of the words? Yes or no? __________ If yes, please give the sentence you used. (If the sentence is not very polite please give it anyway: your name is <u>not</u> on any of the material).

7. Were there any other ways in which you tried to simplify the learning task? Yes or no? __________ If yes, please describe.

8. In studying the list which words did you learn by pronouncing or reading, and which did you learn by spelling out the letters? Please list the words under the appropriate heading.

Words I learned by pronouncing:	Words I learned by spelling out the letters:

9. Did you find the order in which the nonsense words were written on the sheet to be inconvenient? Yes or no? __________ If yes, please say what was wrong with the order, and show how you would have listed them.

Now that you have finished the questionnaire please place EVERYTHING into the envelope and tuck in the flap just as it was when you got it.

Please DO NOT hand in the envelope until asked to do so at the end of the period.

THANK YOU FOR YOUR HELP AND TIME

You may now go on with some school work. Do so quietly so as not to disturb those students who are still doing Part 1 and 2.

APPENDIX G

The S-statistic

THE S-STATISTIC

Statistical models are usually based on rather rigorous mathematical assumptions. In educational research the data are rarely ideal and often violate these assumptions to a certain degree. In the case of the F-test if the data do not meet one or more of the assumptions, it may be difficult to decide whether a significant F is due to a real difference between treatments or whether it is due to the departure from the assumptions.

For the present study, departure from the assumption of homogeneity of variance was of particular importance and it was felt that the critical F-values should be checked. In his paper Bay (May, 1970) describes a computer program which simulates certain statistical models under the conditions of interest to the experimenter. Once the simulated empirical distribution of the F-test, denoted by S-ratio, is obtained from the data in question it may be used as a criterion for testing statistical hypotheses (Bay, 1970, p. 18).

In the present study the variance matrices were the input data and the critical S-values were obtained from the computer simulation; however, in all cases the difference between the critical S-value and the critical F-value was not large enough to change any of the statistical decisions. Consequently, only the F-values are quoted throughout the study.

APPENDIX H

One-Way Analysis of Variance: Schools

ONE-WAY ANALYSIS OF VARIANCE: SCHOOLS

The Subjects

An equal number of male and female Ss was selected from each school according to the following table:

School	Number of Ss per treatment	Total Number of Ss
1. (Leduc)	6	48
2. (St. Albert)	8	64
3. (Stony Plain)	4	32
		144

Within this framework the selection was random. The number of Ss per treatment in each case was determined by the cell with the lowest frequency.

The performance measures lacking homogeneity of variance were: Trials to Criterion, Number Correct on Trial #5, and Number Wrong (Omissions) on Trial #5.

Trials to Criterion

SUMMARY OF ANALYSIS OF VARIANCE

Source of Variation	SS	df	MS	F	p
Schools	30.63	2	15.32	1.70	0.19
Error	1270.81	141	9.01		

Homogeneity of Variance: χ^2 is 7.63, $p< 0.05$

In this case $F_{.95}(2,141)$ is 3.06 while $S_{.95}(2,141)$ is 3.29. Therefore, with or without the correction an observed F of 1.70 is not significant at 0.05 level.

Number Correct on Trial #5

SUMMARY OF ANALYSIS OF VARIANCE

Source of Variation	SS	df	MS	F	p
Schools	7.78	2	3.89	1.12	0.33
Error	487.78	141	3.46		

Homogeneity of Variance: χ^2 is 12.71, $p < 0.01$

In this case $F_{.95}(2,141)$ is 3.06 while $S_{.95}(2,141)$ is 3.30. Therefore, in either case the observed F of 1.12 is not significant at the 0.05 level.

Number Wrong (Omissions) on Trial #5

SUMMARY OF ANALYSIS OF VARIANCE

Source of Variation	SS	df	MS	F	p
Schools	5.79	2	2.89	1.98	0.14
Error	205.96	141	1.46		

Homogeneity of Variance: χ^2 is 33.98, $p < 0.001$

In this case $F_{.95}(2,141)$ is 3.06 while $S_{.95}(2,141)$ is 3.76. Therefore, with or without the correction an observed F of 1.98 is not significant at the 0.05 level.

In summary, the F values for the three measures are much too small to reach significance at the 0.05 level even under conditions of no homogeneity of variance.

APPENDIX I

Tables

TABLE 24

THE NUMBER OF SUBJECTS REPORTING THE CODING OF NONSENSE SYLLABLES BY PRONUNCIATION (P) OR BY SPELLING (S)[a]

List	Code[b]	VUK	VUT	VUX	KUV	TUV	XUV	ZOH	ZOJ	ZOS	HOZ	JUZ	SOZ
100	P	24	26	21	17	20	9	11	7	11	25	26	20
	S	29	29	32	33	31	40	40	44	41	25	28	32
	U	7	5	7	10	9	11	9	9	8	10	6	8

List	Code[b]	VUK	QIF	VUX	KUV[c]	FIQ	XUV	ZOH	QIJ	ZOS	HOZ[c]	JIQ	SOZ
200	P	20	19	22	19	11	7	12	14	19	24	10	15
	S	28	37	33	31	33	36	36	42	35	29	34	28
	U	12	4	5	10	16	17	12	4	6	7	16	17

[a]The percentage data are given in Table 21, p. 173.

[b]Key: P - pronounced
S - spelled
U - unknown (Ss did not specify code)

[c]For a breakdown of the data into Sex x Order see Table 25, Appendix I.

TABLE 24 (continued)[a]

List	Code[b]	VUK	FUX	HUZ	JIH	KUJ	QOS	VOF	QIJ	XIF	XOQ	ZIV	SOZ
	P	43	47	43	24	16	13	38	13	15	10	37	35
300	S	12	8	12	31	37	40	17	39	38	45	17	19
	U	5	5	5	5	7	7	5	8	7	5	6	6

List	Code[b]	VUK	BOF	CEH[c]	GYX	LAJ[c]	MYP	NID	QIJ	ZOS	QUR	TEH[c]	VAW
	P	42	47	30	14	30	25	49	20	41	21	32	40
400	S	17	10	28	44	26	33	9	37	16	35	26	18
	U	1	3	2	2	4	2	2	3	3	4	2	2

[a]The percentage data are given in Table 21, p. 173.

[b]Key: P - pronounced
S - spelled
U - unknown (<u>Ss</u> did not specify code)

[c]For a breakdown of the data into Sex x Order see Table 25, Appendix I.

TABLE 25

THE NUMBER OF MALE AND FEMALE SUBJECTS REPORTING THE CODING OF NONSENSE SYLLABLES BY PRONUNCIATION (P) OR BY SPELLING (S) FOR EACH LIST x ORDER TREATMENT[a]

List	Order	Sex	Code[b]	VUK	VUT	VUX	KUV	TUV	XUV	ZOH	ZOJ	ZOS	HOZ	JOZ	SOZ
100	S	F	P	5	5	5	6	6	2	3	3	4	6	6	6
			S	9	9	9	8	7	11	12	12	11	7	7	7
			U	1	1	1	1	2	2	0	0	0	2	2	2
		M	P	8	8	7	3	4	1	2	1	3	6	7	7
			S	6	6	7	7	7	9	9	9	8	6	7	7
			U	1	1	1	5	4	5	4	5	4	3	1	1
	R	F	P	5	6	3	3	4	2	3	1	1	7	7	3
			S	8	8	9	10	10	11	9	12	12	6	7	10
			U	2	1	3	2	1	2	3	2	2	2	1	2
		M	P	6	7	6	5	6	4	3	2	3	6	6	4
			S	6	6	7	8	7	9	10	11	10	6	7	8
			U	3	2	2	2	2	2	2	2	2	3	2	3

[a]The percentage data are given in Table 22, p. 176.

[b]Key: P - pronounced
S - spelled
U - unknown (Ss did not specify code)

TABLE 25 (continued)[a]

List	Order	Sex	Code[b]	VUK	QIJ	VUX	KUV	FIQ	XUV	ZOH	QIJ	ZOS	HOZ	JIQ	SOZ
200	S	F	P	5	4	5	1	0	0	2	2	4	3	0	2
			S	10	11	10	13	14	14	13	13	11	11	13	11
			U	0	0	0	1	1	1	0	0	0	1	2	2
		M	P	6	5	5	2	2	2	5	5	5	3	3	3
			S	7	8	8	8	8	8	9	9	9	7	8	8
			U	2	2	2	5	5	5	1	1	1	5	4	4
	R	F	P	4	6	5	9	4	2	2	4	6	10	4	6
			S	6	9	8	5	6	9	8	11	7	5	6	4
			U	5	0	2	1	5	4	5	0	2	0	5	5
		M	P	5	4	7	7	5	3	3	3	4	8	3	4
			S	5	9	7	5	5	5	6	9	8	6	7	5
			U	5	2	1	3	5	7	6	3	3	1	5	6

[a]The percentage data are given in Table 22, p. 176.

[b]Key: P - pronounced
S - spelled
U - unknown (Ss did not specify code)

TABLE 25 (continued)[a]

List	Order	Sex	Code[b]	VUK	FUX	HUZ	JIH	KUJ	QOS	VOF	QIJ	XIF	XOQ	ZIV	SOZ
300	S	F	P	11	14	13	7	2	4	11	1	3	1	10	10
			S	4	1	2	8	12	10	4	14	12	14	5	5
			U	0	0	0	0	1	1	0	0	0	0	0	0
		M	P	12	13	8	6	5	3	9	3	3	3	9	9
			S	2	1	4	8	8	9	5	8	11	10	4	3
			U	1	1	3	1	2	3	1	4	1	2	2	3
	R	F	P	10	11	11	8	4	3	9	5	5	3	9	8
			S	5	4	4	7	10	12	5	9	9	12	6	7
			U	0	0	0	0	1	0	1	1	1	0	0	0
		M	P	10	9	11	3	5	3	9	4	4	3	9	8
			S	1	2	2	8	7	9	3	8	6	9	2	4
			U	4	4	2	4	3	3	3	3	5	3	4	3

[a]The percentage data are given in Table 22, p. 176.

[b]Key: P - pronounced
S - spelled
U - unknown (Ss did not specify code)

TABLE 25 (continued)[a]

List	Order	Sex	Code[b]	VUK	BOF	CEH	GYX	LAJ	MYP	NID	QIJ	ZOS	QUR	TEH	VAW
400	S	F	P	11	12	9	5	11	7	12	7	11	8	10	11
			S	3	2	5	10	3	8	2	8	3	6	4	3
			U	1	1	1	0	1	0	1	0	1	1	1	1
		M	P	9	11	8	3	5	5	12	6	11	6	8	9
			S	6	4	7	12	10	10	3	9	4	9	7	6
			U	0	0	0	0	0	0	0	0	0	0	0	0
	R	F	P	9	11	4	0	6	5	13	1	8	3	5	9
			S	6	3	10	15	8	10	2	13	6	11	9	5
			U	0	1	1	0	1	0	0	1	1	1	1	1
		M	P	13	13	9	6	8	8	12	6	11	4	9	11
			S	2	1	6	7	5	5	2	7	3	9	6	4
			U	0	1	0	2	2	2	1	2	1	2	0	0

[a]The percentage data are given in Table 22, p. 176.

[c]Key: P - pronounced
S - spelled
U - unknown (Ss did not specify code)

APPENDIX J

Analysis of Variance for Sex x Lists x Order

TABLE 26

SUMMARY OF THREE-WAY ANALYSIS OF VARIANCE FOR TRIALS TO CRITERION

Source of Variation	SS	df	MS	F
A (Sex)	50.42	1	50.42	7.84**
B (Lists)	713.08	3	237.69	36.94*****
C (Order)	96.27	1	96.27	14.96***
AB	10.88	3	3.63	0.56
AC	1.07	1	1.07	0.17
BC	13.63	3	4.54	0.71
ABC	14.90	3	4.97	0.77
Error	1441.33	224	6.43	

** Significant at 0.01 level, $p < 0.006$

*** Significant at 0.001 level, $p < 0.0002$

***** Significant at 0.00001 level

TABLE 27

SUMMARY OF THREE-WAY ANALYSIS OF VARIANCE
FOR MEAN NUMBER CORRECT PER TRIAL

Source of Variance	SS	df	MS	F
A (Sex)	11.01	1	11.01	6.01*
B (Lists)	150.09	3	50.03	27.29*****
C (Order)	9.05	1	9.05	4.94*
AB	4.02	3	1.34	0.73
AC	2.13	1	2.13	1.16
BC	12.15	3	4.05	2.21
ABC	5.79	3	1.93	1.05
Error	410.06	224	1.83	

* Significant at 0.05 level, $p < 0.02$ for A, and $p < 0.03$ for C

***** Significant at 0.00001 level

TABLE 28

SUMMARY OF THREE-WAY ANALYSIS OF VARIANCE FOR MEAN NUMBER WRONG (OVERT ERRORS) PER TRIAL

Source of Variation	SS	df	MS	F
A (Sex)	0.0027	1	0.0027	0.0048
B (Lists)	13.71	3	4.57	8.30****
C (Order)	0.10	1	0.10	0.19
AB	0.33	3	0.11	0.20
AC	0.07	1	0.07	0.12
BC	0.57	3	0.19	0.35
ABC	4.60	3	1.53	2.78*
Error	123.37	224	0.55	

* Significant at 0.05 level

**** Significant at 0.0001 level, $p < 0.00005$

TABLE 29

SUMMARY OF THREE-WAY ANALYSIS OF VARIANCE FOR MEAN NUMBER WRONG (OMISSIONS) PER TRIAL

Source of Variation	SS	df	MS	F
A (Sex)	11.04	1	11.04	10.78**
B (Lists)	76.39	3	25.46	24.08*****
C (Order)	12.02	1	12.02	11.36***
AB	1.32	3	0.45	0.42
AC	2.58	1	2.58	2.44
BC	8.41	3	2.80	2.64*
ABC	2.10	3	0.70	0.66
Error	236.87	224	1.06	

* Significant at 0.05 level

** Significant at 0.01 level, $p < 0.002$

*** Significant at 0.001 level, $p < 0.0009$

***** Significant at 0.00001 level

APPENDIX K

Orthogonal Comparisons

Standard Error (Sample Calculation)

According to Edwards (1960) the standard error of the weighted difference between the means is obtained by multiplying the error mean square of the analysis of variance by the sum of the squares of the coefficients and dividing it by the number of observations for a single mean. In formula form this may be written as follows:

$$\text{standard error} = \sqrt{\frac{s^2}{n} \sum a_i^2}$$

where a_i is the coefficient, n is the number of observations for a single mean, and s^2 is the error mean square of the analysis of variance.

For trials to criterion the error mean square of the analysis of variance was 6.55 (Table 8, p. 141), and the number of Ss per cell was 30. For all comparisons the coefficients were +1 and -1. From the above formula the standard error for trials to criterion was found to be

$$\text{standard error} = \sqrt{\frac{6.55 \times 2}{30}} = 0.66$$

The t-statistic was obtained by dividing each difference between the means by the standard error. The degrees of freedom for the t-statistic were 232, the number associated with s^2 of the analysis of variance. The difference between the means for the 100 list was -1.26 thus the obtained t was $\frac{-1.26}{0.66} = -1.92$. This was not significant at the 0.01 level ($t_{.99} = 2.60$) nor at the 0.05 level ($t_{.95} = 1.97$).

The calculations for all planned orthogonal comparisons were done according to the above procedure and the results are summarized in the following tables.

Planned Orthogonal Comparisons

Trials to Criterion

Standard Error = 0.66

Error Mean Square = 6.55

List	Mean for S Condition	Mean for R Condition	Obtained t
100	2.47	3.73	-1.92
200	2.30	4.33	-3.08**
300	7.00	7.77	-1.16
400	3.53	4.53	-1.51

** Significant at 0.01 level.

Mean Number Correct per Trial

Standard Error = 0.35

Error Mean Square = 1.87

List	Mean for S Condition	Mean for R Condition	Obtained t
100	10.26	10.25	t<1
200	10.52	9.36	3.27**
300	8.26	8.12	t<1
400	9.78	9.53	t<1

** Significant at 0.01 level

Mean Number of Overt Errors per Trial

Standard Error = 0.19

Error Mean Square = 0.55

List	Mean for S Condition	Mean for R Condition	Obtained t
100	0.73	0.52	1.08
200	0.87	0.85	t<1
300	1.23	1.27	t<1
400	0.71	0.72	t<1

Mean Number of Omissions per Trial

Standard Error = 0.27

Error Mean Square = 1.10

List	Mean for S Condition	Mean for R Condition	Obtained t
100	0.86	1.05	t<1
200	0.54	1.64	-4.04***
300	2.26	2.54	-1.04
400	1.44	1.67	t<1

*** Significant at 0.001